Transcriptional regulation of Hox genes during hindbrain development.

Thesis

UNRESTRICTED

Transcriptional regulation of *Hox* genes during hindbrain development

Stefan Wolfgang Tümpel

A thesis submitted in partial fulfillment of the

requirements of the Open University

for the Degree of Doctor of Philosophy

The Stowers Institute for Medical Research

Kansas City, USA

August 2005

DATE OF SUBMISSION 2 SEPTEMBER 2005
DATE OF AWARD 10 FEBRUARY 2006

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

meinen Eltern

Acknowledgements

The research presented in this thesis was conducted at the Stowers Institute for Medical Research in Kansas City and was supported by Boehringer Ingelheim Funds and the Stowers Institute for Medical Research.

I am extremely grateful to my supervisors Dr. Robb Krumlauf and Dr. Leanne Wiedemann for their constant and endless support, advice and encouragement throughout my studentship. I would like to extend my thanks to all members of the Krumlauf laboratory: Heather Marshall for providing patient instruction in the production of transgenic mice, Kieran Pemberton for instructions on chick electroporation, Christof Nolte for showing me how to generate mutant mice, and Paco Cambronero for helpful experimental support and advice. I also would like to thank Tara Alexander, Rachel Chennault, Kristen Correia, Debra Ellies, Laura Galinañes-Garcia and Mark Parrish for many thoughtful discussions. In particular, I would like to thank Rachel, and also Mark, Tara, Kieran, Laura, Christof and Kristen for proof-reading this manuscript. I would also like to thank Leta Zieber, for ensuring everything ran smoothly.

I am indebted to Joan and Ron Conaway and the members of the Conaway lab, in particular Stephanie Kong, Charles Banks, Chieri Sato, Shigeo Sato and Tingting Yao for their patient help with the biochemical experiments. I would like to thank John McCarthy for his patience and instructions on the EMSA. I would like to thank Thomas Kusch for many discussions and experimental advice and Paul Trainor for many valuable discussions.

Finally, I thank my parents, my brothers and my sisters for their moral support and encouragement. I would especially like to thank Natascha for her continuing support.

Abstract

Hox genes play a crucial role in patterning the anteroposterior axis. Therefore, it is important to understand the mechanisms regulating their expression. One way of identifying regulatory regions is to compare related genomic sequences in order to find conserved elements. Previous work in transgenic mice has defined the elements required to recapitulate *Hoxa2* expression in rhombomeres 3 and 5. However, the mechanisms regulating expression of *Hoxa2* in rhombomeres 2 and 4 were unknown. In this study, I demonstrate that a highly conserved region in the intron of *Hoxa2* contains the control elements directing rhombomere 4 expression. Further, I show that the rhombomere 2 enhancer is located in the second exon of *Hoxa2*.

Due to genome-wide duplication events, the number of *Hox* genes increased during vertebrate evolution. In the pufferfish, this led to the presence of two *Hoxa2* genes, *Hoxa2(a)* and *Hoxa2(b)*. The two co-paralogous genes show differential expression. I compared the control regions of these two genes and identified subtle changes in their *cis*-regulatory regions. Using chick electroporation, I show that several changes in these elements are responsible for the differential expression.

Hoxb1 has a broad expression pattern in the hindbrain at early stages. This expression becomes restricted to rhombomere 4 during later development. I show that Krox20 binds to a highly conserved repressor region and that removal of this element in transgenic constructs leads to an expansion of reporter expression into rhombomeres 3 and 5. Finally, I show that *Hoxb1* expression in the second branchial arch is repressed by a mammalian-specific repressor element.

This work has shed important insight into the mechanisms and factors that modulte expression of *Hoxa2* and *Hoxb1* during hindbrain development.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

A-P	anterior-posterior
ANT-C	*Antennapedia* complex
ATP	adenosine 5'-triphoshate
BA	branchial arch
BMP	bone morphogenetic protein
bp	base pair
Bx-C	*Bithorax* complex
°C	degrees centigrade
CNS	central nervous system
cpm	counts per minute
CR	conserved region
DNA	deoxyribonucleic acid
dpc	days post coitum
DR	direct repeat
DTT	dithiothreitol
EDTA	diaminoethanetetra-acetic acid
EMSA	electrophoretic mobility shift assay
FGF	fibroblast growth factor
g	gram
h	hour
HH	Hamburger and Hamilton stage
HOM-C	homeotic complex
kb	kilobase
M	molar
mg	milligram
min	minute
ml	milliliter
mM	millimolar
NCC	neural crest cells
N-terminal	amino terminal
NP-40	Nonidet P-40
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline

PCR	polymerase chain reaction
pmol	picomole
r	rhombomere
RA	retinoic acid
RAR	Retinoic acid receptor
RARE	retinoic acid response element
RNA	ribonucleic acid
rpm	revolutions per minute
RXR	retinoid X receptor
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
TGF	transforming growth factor
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
U	Units
UTR	untranslated region
UV	ultraviolet
V	volts
W	watts
%(w/v)	grams per 100ml (weight/volume)
%(v/v)	ml per 100ml (volume/volume)

Chapter 1

Introduction

A major question in biology is how a multicellular organism develops from a single cell through embryonic stages to an adult form. Development is a complex process regulated by precisely controlled differential gene expression. This program of differential gene expression results in the determination of different cell fates at varying stages of development. This leads to the specification of cell identity and regional character, which forms the basic, ordered spatial pattern (pattern formation) of the body plan. Recently, great progress has been made in understanding how differential gene activity is regulated in distinct cell populations. Significantly, the *Hox* genes have been shown to be major components of the genetic network regulating vertebrate development. Here, I will report on my analysis of two *Hox* genes, *Hoxa2* and *Hoxb1*, and the mechanisms regulating their expression during craniofacial development.

1.1 Pattern Formation during Craniofacial Development

1.1.1 The Vertebrate Head

Building the head is a complex process that involves contributions from a number of different tissues at distinct times. Many structures such as bones, muscles, nerves and teeth do not develop simultaneously, but rather are assembled in a precise spatiotemporal manner. This is achieved by the unique integration of tissue-specific and spatiotemporally-regulated patterns of gene expression that correlate with cell fate changes. These underlying molecular mechanisms drive morphogenesis by instructing cells to proliferate, migrate, differentiate or undergo apoptosis.

Craniofacial structures derive from the three germ layers. The ectoderm gives rise to the neural, placodal, oral, and epidermal structures, as well as the cranial neural crest cells (NCC). The cranial mesoderm gives rise to the cranial muscles, a subset of the skeletal ele-

ments, and the vasculature. The endoderm gives rise to pharyngeal structures, which will generate early embryonic sensory and cranial neurogenic placodes.

Branchial arches (BA) are composed of paraxial mesoderm and NCC, with an outer covering of ectoderm and an inner core of endoderm (Noden, 1988). The ectoderm generates the epidermis and the sensory neurons of the epibranchial ganglia (Couly and Le Douarin, 1990), while the endoderm forms the taste buds, thyroid, parathyroid, and thymus. It has also been shown that the endoderm induces formation of the placodes (Begbie et al., 1999). The placodes are thickenings of the ectoderm that produce the neuroplasts, forming the distal cranial ganglia. Major components of craniofacial patterning are the NCC. They form the skeletal and connective tissues of the arches, and the mesoderm will form muscle and endothelial cells (Noden, 1983; Schilling and Kimmel, 1994; Trainor et al., 1994).

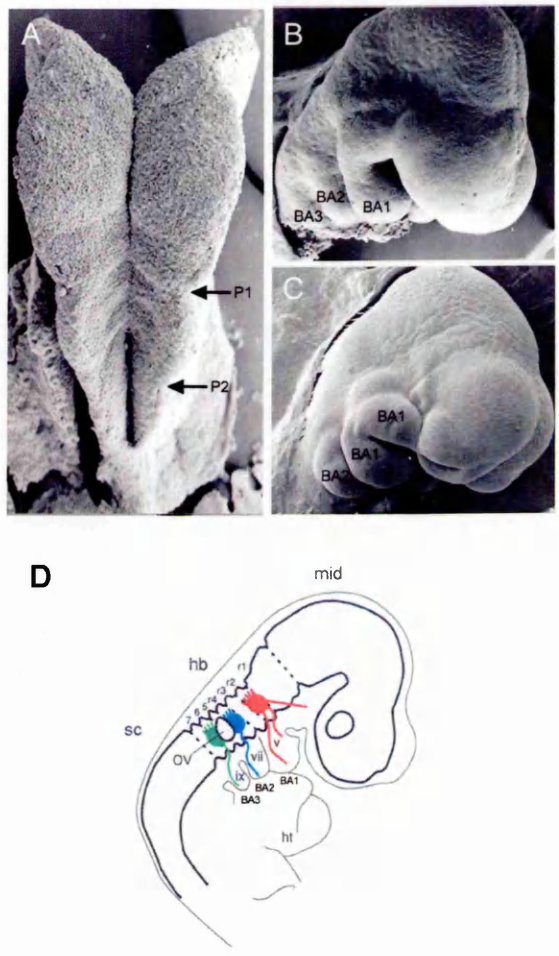


Figure 1.1. Segmentation of the hindbrain.
(A-C) Scanning electron micrograph (SEM) of mouse embryo showing the embryonic head structures at different stages. Figure **A** shows the dorsal view of the anterior structures of an 8.0dpc mouse embryo. The rhombomeres are visible as bulges. The arrows mark the location of the preotic (P1) and postotic (P2) sulci. SEMs of a 9.5dpc **(B)** and of a 10.5dpc **(C)** mouse embryo from the lateral view of the head. The neural tube is closed and the branchial arches (BA1-3) are formed ventral to the hindbrain. **(D)** Schematic diagram of the anterior region of a 9.5dpc embryo showing the location of the rhombomeres (r1-7), cranial nerves (Vth, VIIth and IXth) and branchial arches (BA1-3). BA1, first branchial arch; BA2, second branchial arch; BA3, third branchial arch; ov, otic vesicle; r1-7 rhombomeres 1-7; mid, midbrain; hb, hindbrain; sc, spinal cord; ht, heart; v, trigeminal nerve and ganglion; vii, facial nerve and ganglion; ix glossopharyngeal nerve and ganglion. Photographs courtesy of Liz Hirst (NIMR, London).

The BAs are temporal, metameric structures located on the lateral surface of the head (Figure 1.1). They develop in an anteriorposterior sequence. In mouse, there are six pairs of BAs, with the first and most anterior forming at around 8.0 days post coitum (dpc). Each BA contains a cartilaginous skeletal element and a muscle anlage, which is innervated by a bran-

chial arch-specific cranial nerve and an arch artery. In jawed fish, the BA forming the pharyngeal cavity give rise to structures used for feeding as well as for respiration. The pharyngeal arches of mammals also participate in both of these functions; however, other pharyngeal arch derivatives have adopted new functions. The second arch gives rise to the ear bones. The fourth and sixth arches contribute to the tongue and larynx and, thus, are important for communication.

1.1.2 The Neural Crest Cells

NCC arise at the dorsolateral edges of the neural plate, at the junction of the non-neural ectoderm (surface ectoderm) and the neural plate (neuroepithelium). NCC migrate along both rostro- and ventrolateral pathways below the surface ectoderm and through the mesoderm. The NCC give rise to a variety of tissues. In the cranial region, they form cartilage, bone, connective tissue, neurons and glia of cranial ganglia, and in the trunk region they give rise to pigment cells, sensory neurons, glia, sympathodrenal and Schwann cells. Thus, unlike trunk neural crest, cranial crest cells contribute to both neurogenic and mesenchymal derivatives.

Members of the *Wnt* gene family, that encode for secreted proteins, have been implicated in NCC induction (Garcia-Castro et al., 2002; Saint-Jeannet et al., 1997). It has been shown that the addition of Wnt is sufficient to induce NCC from neural plate tissue. Further, RNA or DNA injection of Wnts into *Xenopus* embryos results in an increase of NCC in whole embryos or in animal cap assays. During induction, NCC undergo an epithelial-to-mesenchymal transition. Subsequently, they migrate away from the neural tube and delaminate (Bronner-Fraser, 2004). This is achieved by downregulation of certain adhesion molecules, such as NCAM, N-cadherin, E-cadherin and cadherin6B. It has been shown that the zinc-finger transcription factors Snail and Slug, which are expressed in pre- and postmigratory NCC, are able to repress *E-cadherin* expression (Bolos et al., 2003; Cano et al., 2000).

Previously it was thought that NCC acquired the patterning information necessary to generate regional differences in craniofacial structures before they migrated away from the

closing neural folds. This was first indicated in Noden's classical grafting experiment where NCC from the first mandibular arch were grafted to a more posterior region, replacing either the NCC of the second (hyoid) or third (visceral) arch (Noden, 1983). The transplanted NCC developed according to their original position, so that first arch derivatives formed in place of second or third arch structures. These results were the basis of the 'pre-program model' of NCC.

More recently, this view has been challenged by a number of experiments in mouse (Golding et al., 2000; Trainor et al., 2002) and zebrafish (Schilling, 2001). It was shown by heterotopic transplantation experiments that cranial NCC differentiate and contribute to structures appropriate to their new axial position. Furthermore, it was shown that NCC transplanted into a new environment do not maintain their normal *Hox* expression profile. The idea, that the environment within the branchial arch plays an important role in NCC patterning, is also supported by experiments showing that *Hox* expression is maintained when NCC are co-transplanted with their neighboring mesodermal cells (Trainor and Krumlauf, 2000). These experiments argue that a small group of NCC can adapt to a new environment and that the mesoderm plays an important role in providing instructive information for *Hox* gene expression, and, therefore, in establishing the identity of NCC.

1.1.3 Development of the Central Nervous System

The tissue that will give rise to the central nervous system (CNS) arises from the neural plate, a group of cells originating from the ectodermal germ layer. It was shown by Hans Spemann and Hilde Mangold that differentiation of the neural plate in *Xenopus* depends on signals from the dorsal lip of the blastopore, the so-called organizer region (Spemann and Mangold, 2001). When the dorsal lip is grafted to the region of a host embryo that normally gives rise to the ventral epidermis, a duplicate body axis is formed.

The secreted growth factor BMP-4, a member of the transforming growth factor β (TGFβ) family, plays an important role in ectodermal cell fate decisions (Piccolo et al., 1996;

Zimmerman et al., 1996). In *Xenopus*, it has been shown that cells expressing a dominant-negative version of the BMP receptor, which blocks BMP signaling, differentiate into neural tissue. This suggests that blocking BMP signaling is sufficient to induce neural differentiation. This observation was further supporter by misexpression experiments of secreted BMP antagonists such as Noggin or Chordin, which were sufficient to mimic the neuralizing capacity of the dorsal lip.

During the process of neurulation, neural cells invaginate to form the neural tube, which can be subdivided into the forebrain (prosencephalon), midbrain (mesencephalon), hindbrain (rhombencephalon), and spinal cord. The neural tube extends from the head to the tail and will give rise to the entire CNS.

1.1.4 Segmentation in the Hindbrain

The vertebrate embryonic hindbrain is segmented into seven cell lineage-restricted compartments, termed rhombomeres (r) (see Figures 1.1A, D and Fraser et al., 1990; Lumsden and Keynes, 1989; Lumsden and Krumlauf, 1996). The most posterior compartment is located at the axial level of the first somite. An eighth putative rhombomere has been identified in the region extending from somite 2 through 5; since it has intermediate characteristics, it has been termed "pseudorhombomeric" (Cambronero and Puelles, 2000). Rhombomeric boundaries form in a precise spatiotemporal manner (Guthrie and Lumsden, 1991). They are generated transiently between 8.0 and 9.5 dpc during mouse development and between Hamburger and Hamilton (HH) stages 6 and 12 during chick development (Figures 1.1).

The hindbrain will give rise to well-defined regions of the brain, including the medulla, pons, and cerebellum (Marin and Puelles, 1995). In addition, the hindbrain provides the organizational plan for craniofacial and CNS structures in the branchial regions. Its segmental organization is critical for patterning and establishing the migratory pathways of neurons, cranial motor nerves and NCC (Lumsden and Keynes, 1989).

The reticular neurons are the earliest neurons to develop in the hindbrain, forming first in r4 and then in r2 and r6 (Sechrist and Bronner-Fraser, 1991). The cranial nerves provide the somatic, visceral sensory and motor innervations for the head and are numbered I-XII (Figure 1.1D). The motor neurons can be subdivided into the branchiomotor and somatic motor neurons. The dorsal region of the hindbrain gives rise to sensory interneurons and relay neurons, whereas the ventral region contains both interneurons and motor neurons. The cranial nerve exit points are located in the even-numbered rhombomeres; the trigeminal (Vth) nerve root exits from r2, the facial (VIIth) nerve root exits from r4 and the glossopharyngeal (IXth) nerve root exits from r6. At later stages, the reticular neurons develop in the odd-numbered rhombomeres (Lumsden, 2004). Thereafter, the neurons in odd-numbered rhombomeres arise, and each of them forms a cluster with the motor neurons of the anteriorly-adjacent rhombomere (Lumsden and Keynes, 1989). The axons of the branchiomotor cranial nerves exit the hindbrain through specialized conduits in the alar plate, which are also the entry points of sensory axons. Later in development, the motor nuclei are organized in longitudinal columns at characteristic dorso-ventral positions. Only r1 does not contain any branchial neurons.

1.1.5 NCC in the hindbrain region

The segmentation of the hindbrain and corresponding rhombomere-specific gene expression are thought to pattern NCC identity and migratory behaviors (Hunt et al., 1991; Köntges and Lumsden, 1996; Lumsden et al., 1991; Schilling and Kimmel, 1994). The migration of NCC has been extensively studied using DiI tracing to follow the cell movements (Kulesa et al., 2000; Kulesa and Fraser, 2000). NCC in the hindbrain migrate into the branchial arches in three stereotypical, major streams. The first arch crest stream (mandibular) arises from the posterior midbrain and anterior hindbrain segments, r1 and r2, and gives rise to the jaw (Lumsden et al., 1991). The second NCC stream (hyoid), which fills the second arch and generates the stapes, the styloid bone, the stylohyoid ligament, the lesser horn of the hyoid bone and connective tissues, migrates primarily from r4. The third stream (postotic), originates

from the posterior region of the hindbrain, r6 and r7, and fills the more caudal branchial arches that give rise to the throat.

Considerably fewer NCC migrate from r3 and r5 to reach the branchial arches, with evidence of cell death (Ellies et al., 2000; Graham et al., 1993; Sechrist et al., 1993). These cells take different anterior and posterior migratory routes. Isolated explants from r3 and r5 are capable of producing NCC when cultured *in vitro*. However, when they are cultured next to even-numbered rhombomeres, no NCC are produced (Graham et al., 1993; Graham et al., 1994). Two key regulators of this phenomenon are the signaling molecule BMP-4 and the homeobox transcription factor Msx2 (Graham et al., 1994). Both *BMP-4* and *Msx2* are expressed in r3 and r5 and it has been shown that this expression is dependent upon signals from the adjacent rhombomeres. If recombinant BMP-4 protein is added to cultured r3 or r5 explants, the expression of *BMP-4* and *Msx2* is maintained, and neural crest migration is inhibited. Reinvestigation of this process has shown that the absence of NCC migration from r3 and r5 is the result of positional cues read from the environment, rather than from regional cell death prior to migration of the NCC (Farlie et al., 1999). More recently, it was shown that overexpression of the Wnt antagonist sFRP2 inhibits BMP signaling and blocks cell death (Ellies et al., 2000). This type of feedback loop between Wnt and BMP signaling may function to prevent the formation of needless muscle attachment sites (Ellies et al., 2002). Altogether, these data suggest that signals from the microenvironment adjacent to the neural tube modulate NCC migration and that interaction between the Wnt and BMP signaling pathways play an important role in eliminating NCC before migration.

1.1.6 Rhombomere Border Formation

Segmentation is a developmental mechanism by which compartments are established. The segregated blocks of cells have distinct characteristics (Lawrence and Struhl, 1996). In the hindbrain, the segmented compartments are called rhombomeres. It has been shown by cell lineage tracing studies in chick that neuroepithelial cells can move within the boundaries of a

single rhombomere, but not from one rhombomere to another (Fraser et al., 1990). Furthermore, *in vitro* studies have shown that, following mixing, even- and odd-numbered rhombomeric cells will segregate away from each other and form discrete patches and stripes. This behavior is not observed when either odd-numbered or even-numbered rhombomeric cells are mixed together (Wizenmann and Lumsden, 1997). This data suggest that even- and odd-numbered rhombomeric cells display different affinities towards each other.

Possible candidate mediators for these differential affinities are members of the receptor tyrosine kinase family, the Ephs, and their membrane-bound ligands, the ephrins. The receptors *EphA4, EphB2*, and *EphB3* are expressed in odd-numbered rhombomeres, whereas the ligands, *ephrins-B1, -B2* and *–B3*, are expressed in the even-numbered rhombomeres (Xu et al., 2000). This complementary pattern of expression implies that interactions between the receptors and ligands occur at the interface of adjacent rhombomeres. The fact that these molecules play a major role in establishing boundaries between rhombomeres has been demonstrated in zebrafish and *Xenopus* (Xu et al., 1995; Xu et al., 1999). Mosaic activation of an ephrin ligand in the hindbrain of zebrafish embryos, results in the cells of even-numbered rhombomeres migrating away from their non-expressing neighbors to the rhombomeric boundaries. In addition to this, there was a decrease in the number of cells left in the even-numbered rhombomeres. In contrast, mosaic activation of Eph receptors results in cells sorting of the boundaries to the odd-numbered rhombomeres. This confirms that the interplay of Eph/ephrin provides a mechanism of sharpening and maintaining the boundaries between rhombomeres.

1.2 The *Hox* Genes

Hox genes belong to a large family of transcription factors called the homeobox gene family. Each member contains a similar DNA-binding region of 60 amino acids called the homeodomain (Gehring et al., 1990; Otting et al., 1988). This domain is encoded by a 180 bp DNA sequence termed the homeobox (McGinnis et al., 1984). *Hox* genes were originally identified in the fruitfly *Drosophila melanogaster* (Lewis, 1978) and were subsequently identified

in all metazoans, suggesting that they play an important, evolutionarily-conserved role during development (Carroll, 1995; Finnerty et al., 2004; Lee et al., 2003; McGinnis and Krumlauf, 1992; Takio et al., 2004). Using a variety of model systems, including *Drosophila*, fish, frog, chick, and mouse, *Hox* genes have been shown to establish axial identity in developing embryos (McGinnis and Krumlauf, 1992).

1.2.1 Homeotic mutations in *Drosophila*

Mutations that result in altered identity of adult segments and appendages have been documented in *Drosophila* (Lewis, 1978). This phenomenon, where one body part of an organism has its identity changed to that of another, is called homeotic transformation. These Hox mutations were identified in two regions of the *Drosophila* genome. The region in which mutations affect the more anterior body parts was termed the *Antennapedia* complex (ANTP-C), and the region in which mutations affect the more posterior body parts was termed the *bithorax* complex (BX-C). For example, the homeotic mutant bithorax has two second thoracic segments instead of a second and a third thoracic segment, because the third thoracic segment has been transformed to resemble a second thoracic segment. This results in a fly with four wings. The *ANTP-C* consists of five *Hox* genes, and the *BX-C* includes three *Hox* genes. *Hox* genes are expressed in overlapping regions along the anteriorposterior (A-P) axis (Akam, 1987). In this way, each segment retains a unique combination of *Hox* gene expression, which generates its unique segmental identity.

1.2.2 The vertebrate *Hox* genes

Similar complexes of homeotic genes have been identified in other animals (Figure 1.2). There are four unlinked *Hox* complexes, *Hoxa, Hoxb, Hoxc* and *Hoxd*, containing a total of 39 genes, in chicken, mouse and human (Figure 1.2). The vertebrate *Hox* genes can be subdivided into 14 paralogous groups, with each group having homology to its equivalent *Hox* member in *Drosophila* (Figure 1.2). For example, members of the first paralogous group include *Hoxa1*, *Hoxb1* and *Hoxd1*.

The homeotic genes are related not only by sequence similarity, but also by their clustered organization. Unlike in *Drosophila*, the genes in the vertebrate *Hox* cluster are aligned in the same 5' to 3' transcriptional orientation (McGinnis and Krumlauf, 1992). Comparison of the DNA and protein sequences of the four clusters revealed that they are homologous and appear to have evolved through duplication from an ancestral cluster that contained representatives of all members of the family (Kappen et al., 1989; Zhang and Nei, 1996). As each of the four complexes do not contain all 14 paralogous group members, it is likely that a number of them have been lost during evolution.

In invertebrates, only one *Hox* cluster has been identified, with *Drosophila* having 8 genes and Amphioxus having 14 genes (Ferrier et al., 2000; Minguillon et al., 2005). This cluster is thought to have been duplicated to four clusters (A-D in amniotes) on different chromosomes via two rounds of genome duplications (Ohno, 1993). In teleosts (e.g. pufferfish, zebrafish), further duplication occurred during evolution, leading to 7 clusters (Amores et al., 2004; Prince et al., 1998). The pufferfish has duplicate copies of *Hox* clusters that are present in a single copy in amniotes, which includes the *Hoxa*, *Hoxb* and *Hoxd* clusters. Interestingly, in zebrafish, at least 11 of these duplicated genes have been retained (Prince and Pickett, 2002). For some of these duplicated genes

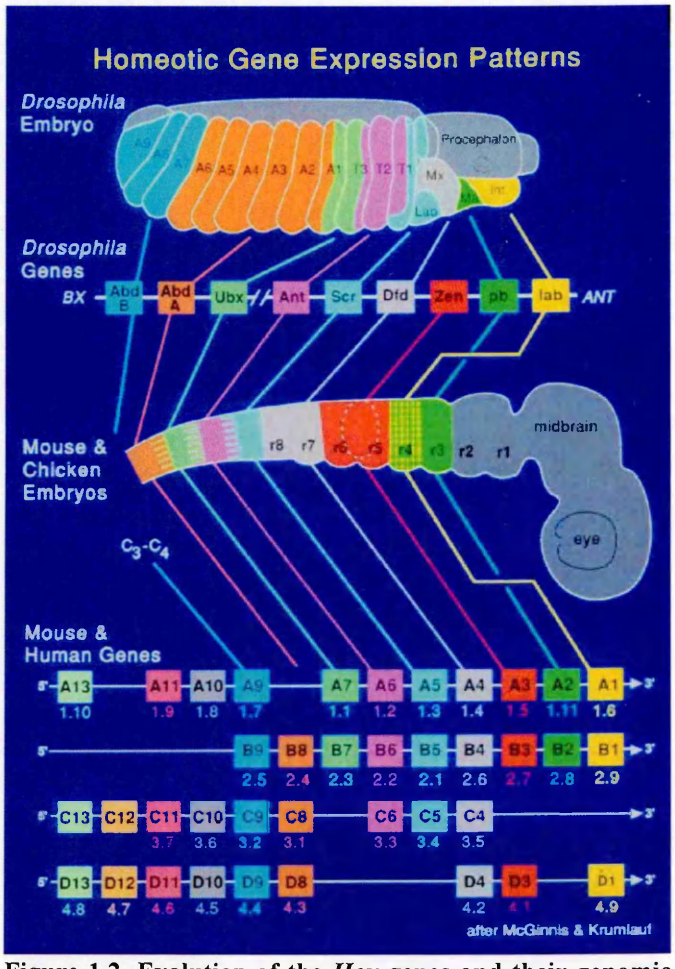


Figure 1.2. Evolution of the *Hox* genes and their genomic organization.
Schematic diagram shows *Hox* gene expression in *Drosophila* and in vertebrates which reflects their organization in the genome. There is one *Hox* cluster in *Drosophila* and four separated complex in most vertebrates (e.g. mouse, chicken and human), which are termed *Hoxa, Hoxb, Hoxc* and *Hoxd*. The vertical order of the vertebrate *Hox* genes shows that they share high sequence similarities and that they arose by duplications and divergence. Members of genes ordered in one vertical row are termed a paralogous group, which extent from 1 to 13.

their expression pattern differs, which has been correlated with differences in their regulatory regions (McClintock et al., 2001; McClintock et al., 2002). This implies that these genes have been retained during evolution as a result of sub-functionalization.

Hox genes govern the patterning of the CNS (Schneider-Maunoury et al., 1993; Studer et al., 1996), the limbs (Dollé et al., 1989a; Izpisua-Belmonte et al., 1991b; Morgan et al., 1992), vertebrae (Bailey et al., 1997; Chisaka and Capecchi, 1991; Kessel and Gruss, 1990; Le Mouellic et al., 1992; Lufkin et al., 1991) and craniofacial structures (Balling et al., 1989; Chisaka and Capecchi, 1991; Lufkin et al., 1991) by providing positional molecular instructions, or a *Hox* code, along the A-P axis (Hunt et al., 1991; Kessel and Gruss, 1991).

1.2.3 HOX colinearity

Since all of the *Hox* genes are orientated in the same direction, with the only exception of the *Drosophila Deformed* gene, the cluster can be described as having a 3' to 5' polarity. The *ANTP-C* genes (the more 3' genes) are expressed in more anterior regions of the fly while the *BX-C* genes (the more 5' genes) are expressed more posteriorly. This relationship implies that there is a correlation between the order of the genes along the cluster and their expression domains along the A-P axis of the embryo. This property is termed spatial colinearity (Lewis, 1978).

This ordered relationship has also been also observed in vertebrate *Hox* complexes, in that each successive gene, moving in a 3' to 5' direction along the complex, has a more posterior expression boundary along the A-P axis of the embryo. This phenomenon is observed in many different tissues, including nervous system, neural crest, branchial arches, paraxial mesoderm (somites), gut and limbs (Chavrier et al., 1990; Dollé et al., 1989b; Dressler and Gruss, 1989; Graham et al., 1989; Hunt et al., 1991; Izpisua-Belmonte et al., 1991a; Izpisua-Belmonte et al., 1991b; Kessel and Gruss, 1991). In the hindbrain, 3' *Hox* genes have boundaries of expression that map to distinct rhombomere boundaries. The paralogous genes of the *labial* group (group 1, e.g. *Hoxb1*) have expression boundaries that are set more anteriorly in

the hindbrain than members of the *Deformed* group (group 4, e.g. *Hoxb4*). However, this model is not consistent with the more 3' *Hox* genes. For example, a member of the *labial* group, *Hoxb1*, is expressed more posteriorly, in rhombomere 4, than a member of the *Proboscipedia* group, *Hoxa2*, which has its anterior expression limit at the boundary of r1 and r2 (Krumlauf, 1993).

Another feature of the *Hox* gene family is temporal colinearity. *Hox* genes at the 3' end are activated earlier than genes at the 5' end of the clusters. The paralogous genes of the *labial* group (group 1) are expressed earlier during development than those of the *Proboscipedia* group (group 2).

Yet another feature of the *Hox* gene family is quantitative colinearity, which describes the responsiveness of these genes to retinoic acid (RA) and fibroblast growth factor (Fgf) signaling (Bel-Vialar et al., 2002; Papalopulu et al., 1991). The responsiveness to RA has been studied both in cell lines and embryos. *Hox* genes are differentially induced by RA according to their location on the cluster, in a time and concentration dependent manner. *Hox* genes at the 3' end of the cluster are activated earlier during development and show a higher sensitivity to RA induction than the more 5' *Hox* genes (Chambeyron and Bickmore, 2004; Dekker et al., 1992; Oudejans et al., 1990; Papalopulu et al., 1991; Simeone et al., 1989).

Another example of quantitative colinearity seen in the *Hox* genes is the response to Fgf signaling (Bel-Vialar et al., 2002). It was shown that the most 5' *Hox* genes *(Hoxb6-Hoxb9)* can be ectopically induced by Fgf overexpression (Bel-Vialar et al., 2002). On the other hand, this is not the case for the more 3' *Hoxb* genes (*Hoxb1* and *Hoxb3-Hoxb5*) which do not show a response to ectopic Fgf activation (Bel-Vialar et al., 2002). This is in contrast to the RA sensitivity of *Hox* expression, which shows a stronger effect on the 3' *Hox* gene members of the complex.

The mechanism behind colinearity is not fully understood. Three models have been proposed to explain these features of *Hox* genes (Kmita and Duboule, 2003). The first mechanism is based on chromatin remodeling, in which repressive or silencing factors could account

for the transcriptional activity of the *Hox* genes (Gould et al., 1997). In this model, the more 3' genes in the *Hox* cluster reside in a decondensated domain. Over time, the more 5' *Hox* genes are chromatin-dependently accessible to the transcription machinery. Support for this model comes from a study where ES cells were used to follow the temporal activation of *Hox* genes by monitoring histone modification and chromatin condensation (Chambeyron and Bickmore, 2004).

The second model describes that the spatial and temporal colinearity can be explained by a local *cis*-regulatory region, which shows differential affinities to upstream signals (e.g. RA or Fgf) throughout the clusters (Bel-Vialar et al., 2002; Oudejans et al., 1990). Support for this idea arose from a study which shows that adjacent *Hox* genes within the cluster share local enhancer elements, which are activated in the same time window (Sharpe et al., 1998).

The third model proposes that global enhancer elements outside the *Hox* clusters are responsible for the temporal colinearity. Since the global elements are located on either side of the *Hox* genes, they have asymmetric distance to the *Hox* genes and, therefore, are differentially regulated. Some of these global elements have been identified, and it was shown that they are able to regulate the expression of several genes at different times (Kmita et al., 2002; Spitz et al., 2003; Zakany et al., 2001).

All of these models are supported experimentally, and it is likely that the mechanism underlying colinearity involves a combination of these models.

1.3 Retinoic Acid

1.3.1 Retinoic Acid in development

RA is a derivative of vitamin A (all-*trans*-retinol) and is an essential signaling molecule for early embryogenesis RA plays a critical role in specifying the A-P axis in the limb bud and central nervous system (Marshall et al., 1992; Schneider et al., 2001; Tickle et al., 1982). Also RA has been shown to play a crucial role during hindbrain development (Gavalas, 2002; Maden, 2002).

Embryonic tissue is able to synthesize endogenous RA by metabolism of Vitamin A. In the blood, retinol circulates bound to retinol-binding proteins. Retinol in the cell is enzymatically converted to retinal by either retinol or alcohol dehydrogenase (RoDHs or ADHs), and then to RA by the retinaldehyde dehydrogenases (RALDHs). The RALDHs contain different members, with RALDH-1, -2 and -3 being the significant during embryogenesis (Duester, 2000). There are two isoforms of RA, all-*trans*-RA and 9-*cis*-RA, which are ligands for different nuclear receptors. Cells requiring RA also contain cellular retinoid binding proteins (CRBP-I, CRABP-I and CRABP-II). These proteins are also responsible for generating the RA gradient across the A-P axis (Duester, 2000; Niederreither et al., 1999).

RA interacts with different sets of nuclear retinoic acid receptors (RAR), which are able to bind to retinoic acid response elements (RARE). RAREs consist of two related sequence motifs with the consensus sequence PuG(GT)TCA(X)nPuG (GT)TCA, in which Pu is a purine base and (X)n the number of nucleotides which separate the two motifs (Leid et al., 1992; Mainguy et al., 2003). It has been shown that the spacing between the two motifs functions as a code for different nuclear receptors. Heterodimers composed of RXRs and RARs preferentially bind to direct motifs separated by 2 (DR2) or 5 (DR5) nucleotides. Motifs of DR4, DR3 and DR1 are preferentially bound by thyroid hormone receptors, vitamin receptors and retinoid X receptors (RXR), respectively (Kliewer et al., 1992).

In the CNS, RA is involved in patterning the hindbrain and anterior spinal cord (Gale et al., 1999; Maden, 2002; Marshall et al., 1992). Furthermore, RA has been shown to play an important role in regulating interneuron and motor neuron development (Sockanathan and Jessell, 1998). The effect of excess RA has been studied in a number of different model systems. It has been demonstrated that there is a concentration-dependent effect of excess RA on rhombomere-specific gene expressing (Gale et al., 1999; Godsave et al., 1998; Kessel and Gruss, 1991; Marshall et al., 1992). For, example, it has been shown that RA treatment results in an anterior shift of *Hox* gene expression, and transformation of rhombomeric identities has been described (Marshall et al., 1992). Quail embryos kept on a vitamin A deficient (VAD) diet,

exhibit a loss of posterior rhombomeres including r4, r5, r6 and r7, and expansion of the remaining rhombomeres r1, r2 and r3 (Gale et al., 1999). This observation is further confirmed by studies of chick embryos cultured in an antagonist to all three retinoic acid receptors (Dupe and Lumsden, 2001). Culturing the embryos in a high concentration of this antagonist leads to loss of posterior hindbrain structures and the enlargement of r4, based on the expansion of *Hoxb1* expression. Treating embryos at successively earlier stages leads to a sequential loss of posterior rhombomeres.

The conclusion from these and other experiments is that an excess of RA signaling disrupts development of the anterior hindbrain region, whereas decreased signaling disturbs the more posterior hindbrain region. This led to the formulation of a model to explain the function of RA in the posterior hindbrain (Gavalas, 2002; Maden, 2002). According to this model, RA is released from the posterior region of the hindbrain creating a morphogen gradient that decreases towards the anterior limit of the hindbrain. This gradient, together with the differential expression of the RA-receptors, is thought to be responsible for creating an RA activity gradient (Gavalas, 2002). Increasing concentrations of RA are required for the specification of the more posterior rhombomeres, while specification of the anterior region requires lower RA concentrations.

The functional mechanisms by which RA regulates the expression of genes involved in hindbrain patterning have been studied in great detail. RAREs have been identified in the regulatory regions of several *Hox* genes (Frasch et al., 1995; Gould, 1997; Marshall et al., 1994). Through this mechanism, the gradient of RA and the differential expression of its effectors are translated into the expression of rhombomere-specific segmentation and segment identity genes.

1.3.2 Retinoic acid and its receptors

Two families of nuclear retinoid receptors, involved in regulating developmental genes by RA, have been identified. They are the *retinoic acid receptors* (*RARs*) and the *retinoid X*

receptors (*RXRs*). The *RAR* family is composed of three members, *RARα*, *RARβ* and *RARγ*, and a number of splice variants have been reported for each gene. In mammals, all-*trans* RA and 9-*cis* RA bind to these receptors.

The *RXRs* family consists of three members, *RXRα, RXRβ*, and *RXRγ*, each of which is also processed into multiple isoforms (Mangelsdorf et al., 1990). These receptors are only bound by 9-*cis* RA (Heyman et al., 1992). In the absence of RA, RXRs and RARs bind to the RAREs where they interact with co-repressors such as Nco-R and SMRT. In the presence of RA, the co-repressors are released and, in turn, recruits co-activators like SRC1 and RIP140 (Glass and Rosenfeld, 2000). The expression pattern of the nuclear retinoid receptors is temporally and spatially restricted in various regions, which implies that they have a role in craniofacial development (Ruberte et al., 1991). *RARα* transcripts are first detected during gastrulation in the epiblast and the mesoderm of the primitive streak (Ruberte et al., 1991). At 8.5 dpc, *RARα* is strongly expressed in the forebrain neuroepithelium and in the neuroectoderm of the posterior hindbrain. Expression is also stronger in the NCC migrating from the posterior hindbrain, and also in the NCC which migrate into the frontonasal mesenchyme and first branchial arch. By late 8.5 dpc, *RARα* is ubiquitously expressed throughout the embryo, although strong expression persists in the migrating NCC (Dollé et al., 1990; Ruberte et al., 1991). At 9.0dpc, *RARα* expression is downregulated establishing an anterior limit of expression at the r3/4 boundary of the hindbrain.

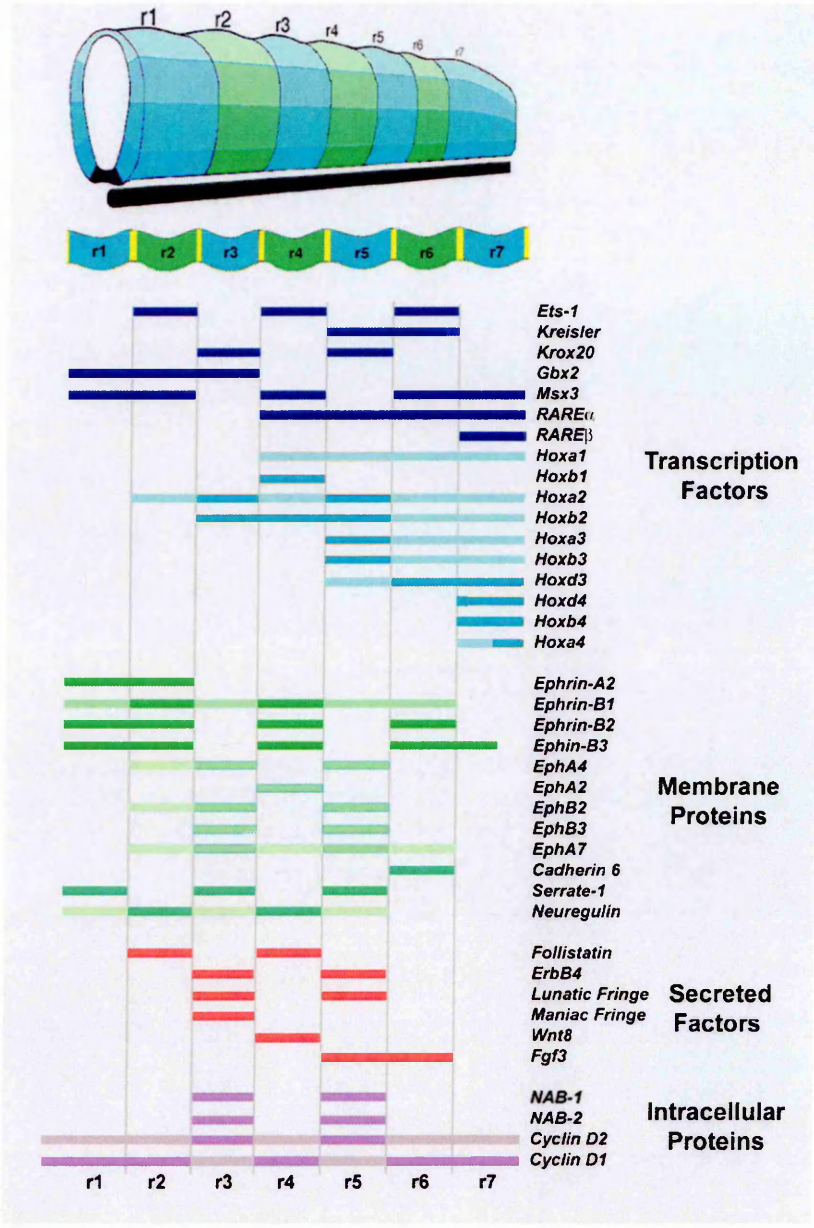


Figure 1.3. Segmented expression of various genes in the hindbrain. Schematic diagram of the hindbrain and examples of genes including transcription factors, membrane proteins, secreted factors and intracellular proteins, which show segment-restricted expression. The intensity of color reflects the level of expression (modified from Lumsden and Krumlauf, 1996).

The expression pattern of *RARβ* during early mouse development has also been described in great detail (Ruberte et al., 1991). *RARβ* transcripts are abundant in the crest-derived frontonasal mesenchyme, but not in the branchial arches. At 7.5 dpc, *RARβ* is initially expressed in the lateral mesoderm extending into anterior regions. From 8.0 dpc, *RARβ* transcripts are detected in the neural epithelium of the posterior hindbrain, which, by 9.5 dpc, have an anterior limit at the level of the first somite. By 10.0 dpc, the anterior limit of *RARβ* expression coincides with the r6/7 boundary in the hindbrain and extends posteriorly along the neural tube (Izpisua-Belmonte et al., 1991a).

1.4 Genetic Control of Segmentation in the Hindbrain

Hindbrain segmentation and specification of rhombomeres require a series of genetic interactions. Several transcription factors are known to play an important role in hindbrain segmentation (Figure 1.3). One of the most important classes of transcription factors are the *Hox* genes. Members of the first four *Hox* paralogous groups are expressed in the hindbrain

where they show rhombomere-specific expression in either one or more rhombomeres (Figure 1.3). The regional identity of the hindbrain is specified, in part, by these *Hox* genes. In addition, other transcription factors, like *kreisler* and *Krox20*, are also important for patterning the hindbrain. Signaling molecules, including RA, and nuclear and membrane receptors such as *EphA4*, are expressed dynamically in the developing hindbrain (see Figure 3 and Lumsden and Krumlauf, 1996).

The following description of the regulatory network in the hindbrain and in the BAs is subdivided into four different time periods. The temporal description of the gene expression is based on mouse data, although results from other model systems are included.

1.4.1 Early Genetic Network (7.5 - 8.0 dpc)

At 7.5 dpc, the hindbrain has not yet developed distinct morphological regions; however, molecular domains have already been established (Figure 1.4). *Hoxa1* and *Hoxb1* are the earliest *Hox* genes expressed in the hindbrain and they both show an anterior limit of expression at the presumptive (pre-) r3/r4 border. This anterior limit of expression exactly coincides for these two genes (Barrow et al., 2000; Frohman et al., 1990). The expression of both genes is initiated by RA, mediated through highly conserved RARE located in their regulatory regions (see Figures 1.4, 1.5A,B and Frasch et al., 1995; Langston and Gudas, 1992; Marshall et al., 1994). The first RARE of *Hoxb1* that is activated following RA binding is located 3' of the *Hoxb1* coding region and regulates expression in neural and mesodermal tissues, and in the primitive streak (Figure 1.5B). This RARE belongs to the DR2 type. Another enhancer is located further upstream, which regulates expression of *Hoxb1* in the somites, node, and lateral mesoderm (see Figure 5B Marshall et al., 1994). The precise regulatory mechanism directing *Hoxb1* expression in these tissues is currently unknown. Another RARE element, located 5' of the coding region of *Hoxb1*, functions as an r3/5 repressor element (Studer et al., 1994). This element restricts expression of *Hoxb1* to r4 (Figures 1.4 and 1.5B). The removal of this element results in expansion of *Hoxb1* expression into r3 and r5 (Studer et al., 1994).

The single RARE of *Hoxa1*, located 5' of the gene, is a DR5 RARE and regulates expression in the notochord, floor plate, caudal neural tube, and gut epithelium (Figure 1.5A and Frasch et al., 1995). Adjacent to this RARE, a highly conserved region (CE2) has been identified, that is sufficient to regulate early somitic and mesenchymal *Hoxa1* expression (Figure

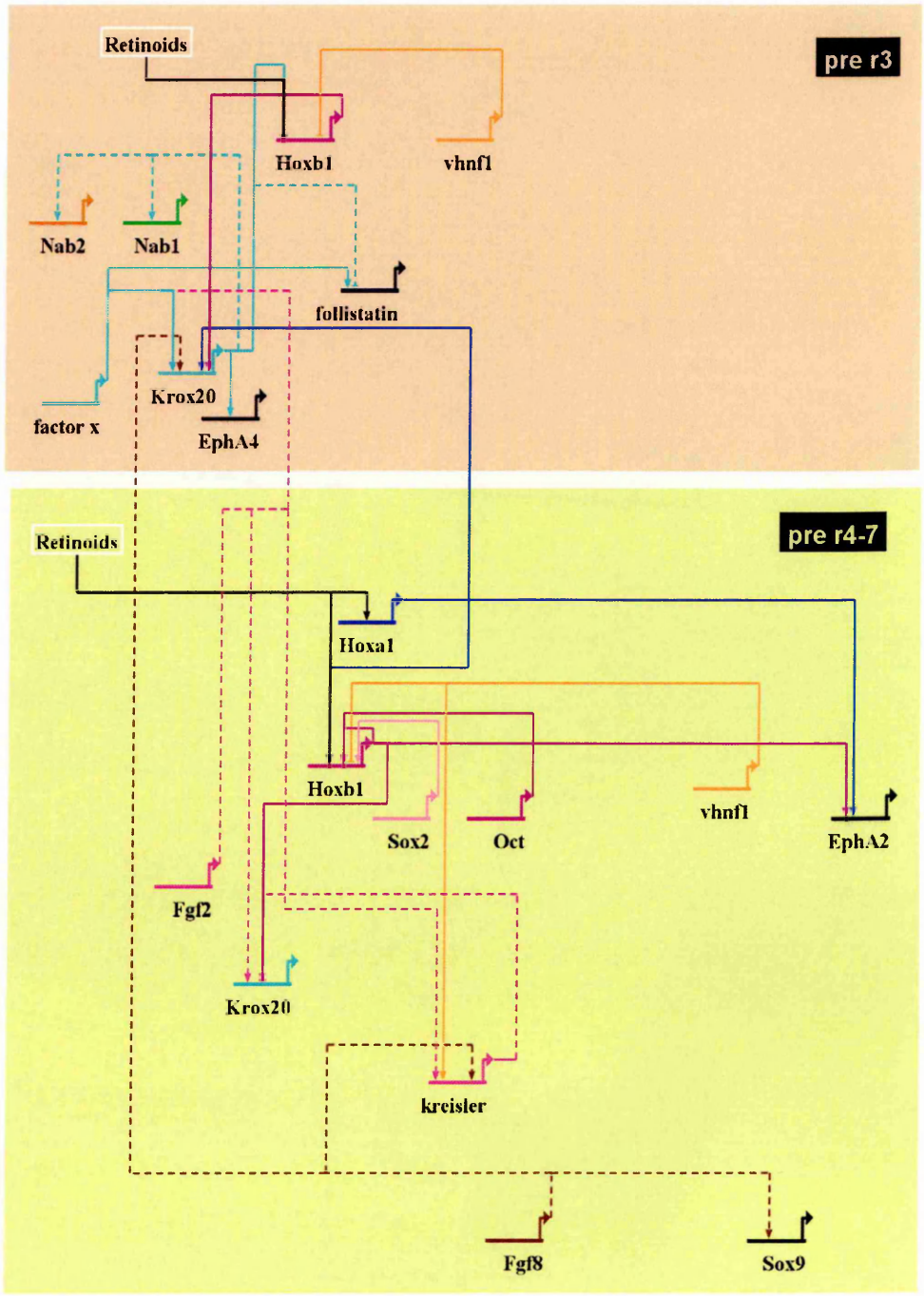


Figure 1.4. Genetic regulatory networks in the hindbrain region at 8.0dpc.
At this stage the hindbrain can be subdivided into two regions, the presumptive rhombomere 3 (pre r3) regions and presumptive rhombomere 4 to 7 (pre r4-7) region. Dotted line illustrates that the interaction between the genes have been not been shown to be direct. Arrow head at the end of the line represents activation, whereas perpendicular line at the end of the line illustrates repression between the connected genes.

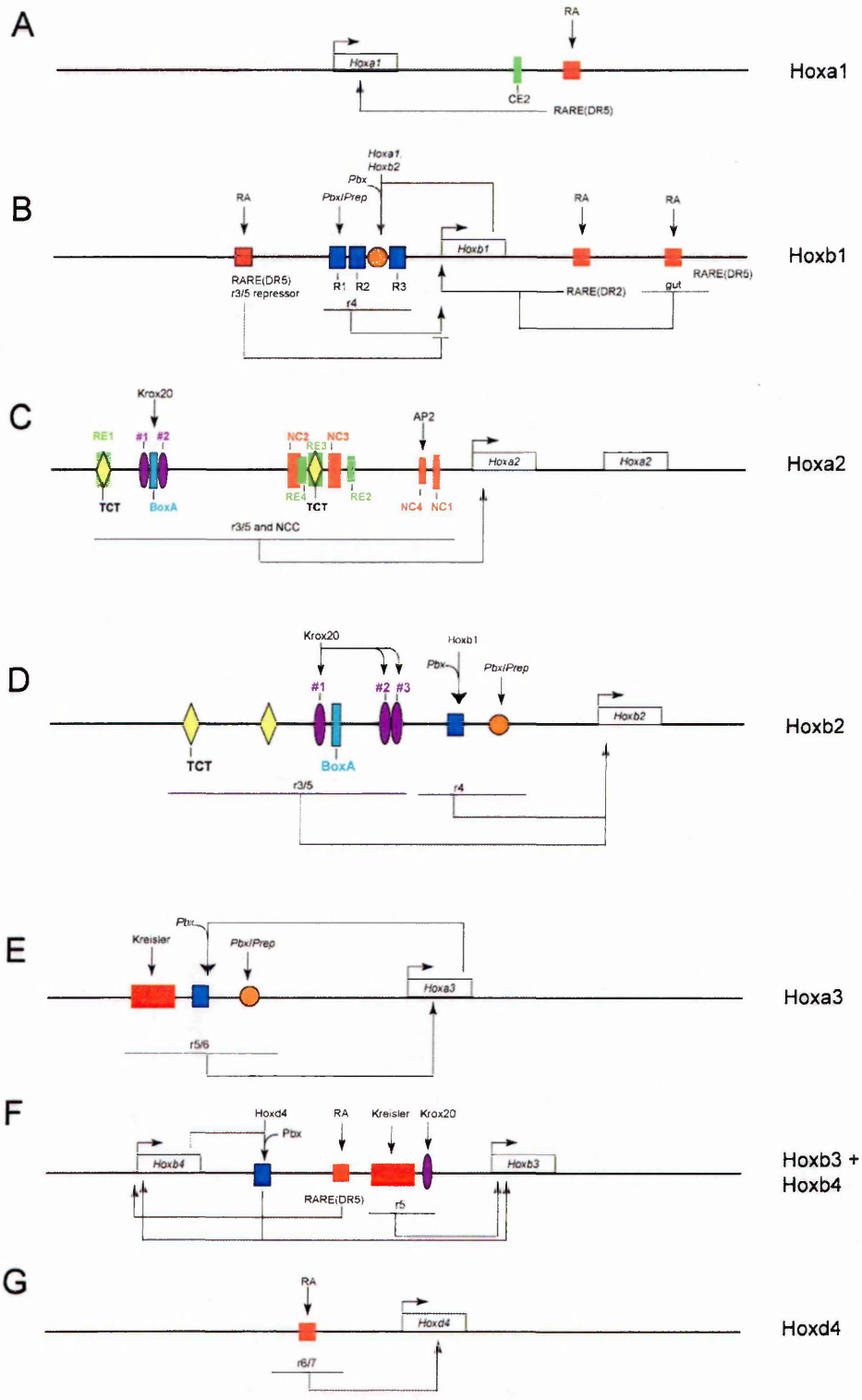


Figure 1.5. Schematic diagrams illustrating the regulatory modules of various *Hox* genes.
On the top of the module the *trans*-acting factors are listed, on the bottom of the module the embryonic expression domains for each of the modules are listed. See text for more details.

1.5A, Thompson et al., 1998). In mouse and chicken, *Hoxd1*, the third group 1 member, is not expressed in the CNS (Frohman and Martin, 1992; Hunt et al., 1991).

During early development (0-3 somite) *Hoxb1* expression in pre-r4 becomes increasingly stronger. An additional molecular domain is established in the hindbrain zinc-finger

transcription factor Krox20. Early expression in the pre-r3 region expands during early development (0-3s) (Figure 1.4 and see also Irving et al., 1996). The anterior limit of *Krox20* expression corresponds to the position of the preotic sulcus, which marks the r2/3 boundary. This early expression of *Krox20* may be initiated by members of the fibroblast growth factor (Fgf) family (Marin and Charnay, 2000; Walshe et al., 2002). It has been shown that Fgf-2, -4, and -8 can ectopically induce *Krox20* expression in the hindbrain of chick embryos (Figure 1.4). Experiments in which *Fgf* transcripts have been knocked-down in zebrafish embryos support this idea (Walshe et al., 2002). Furthermore, the early expression domain of *Krox20* in compound mutants for Hoxa1 and Hoxb1 is shifted more posteriorly suggesting a role for these *Hox* genes in repressing *Krox20* expression in the pre-r4 domain (Barrow et al., 2000).

At this stage, transcripts of the tyrosine receptor kinase *EphA4* are also detectable in the pre-r3 region (Nieto et al., 1992). This early expression is directly triggered by Krox20 as shown by mouse transgenic analysis (Figure 1.6 and Theil et al., 1998). At this stage, *Nab1* and *Nab2* also appear at this early stage in pre-r3. It has been shown by genetic studies that their expression is also dependent on Krox20 (Mechta-Grigoriou et al., 2000). Nab proteins have been shown to repress transcriptional activation, mediated by Krox20 and another zinc-finger transcription factor, NGFI-A, by turning the transcriptional activator into a transcriptional repressor. (Russo et al., 1995; Svaren et al., 1996). Another early hindbrain marker is follistatin, which is expressed at 8.0 dpc from the pre-r1/2 boundary through pre-r4, leaving a gap in pre-r3 (Albano et al., 1994). The factors that regulate follistatin expression are unknown. However, it has been shown that Krox20 downregulates the expression domain in r3 (Seitanidou et al., 1997).

Additionally, in the more posterior region of the hindbrain, *kreisler* is expressed in the presumptive region of r5 and r6. *kreisler* expression is initiated at 7.5 dpc in the pre-r5 region and is later expressed in r5 and r6 until 9.0 dpc, when it is quickly downregulated in both rhombomeres. Further analysis has shown that *kreisler* expression also depends on Fgf signaling (Marin and Charnay, 2000).

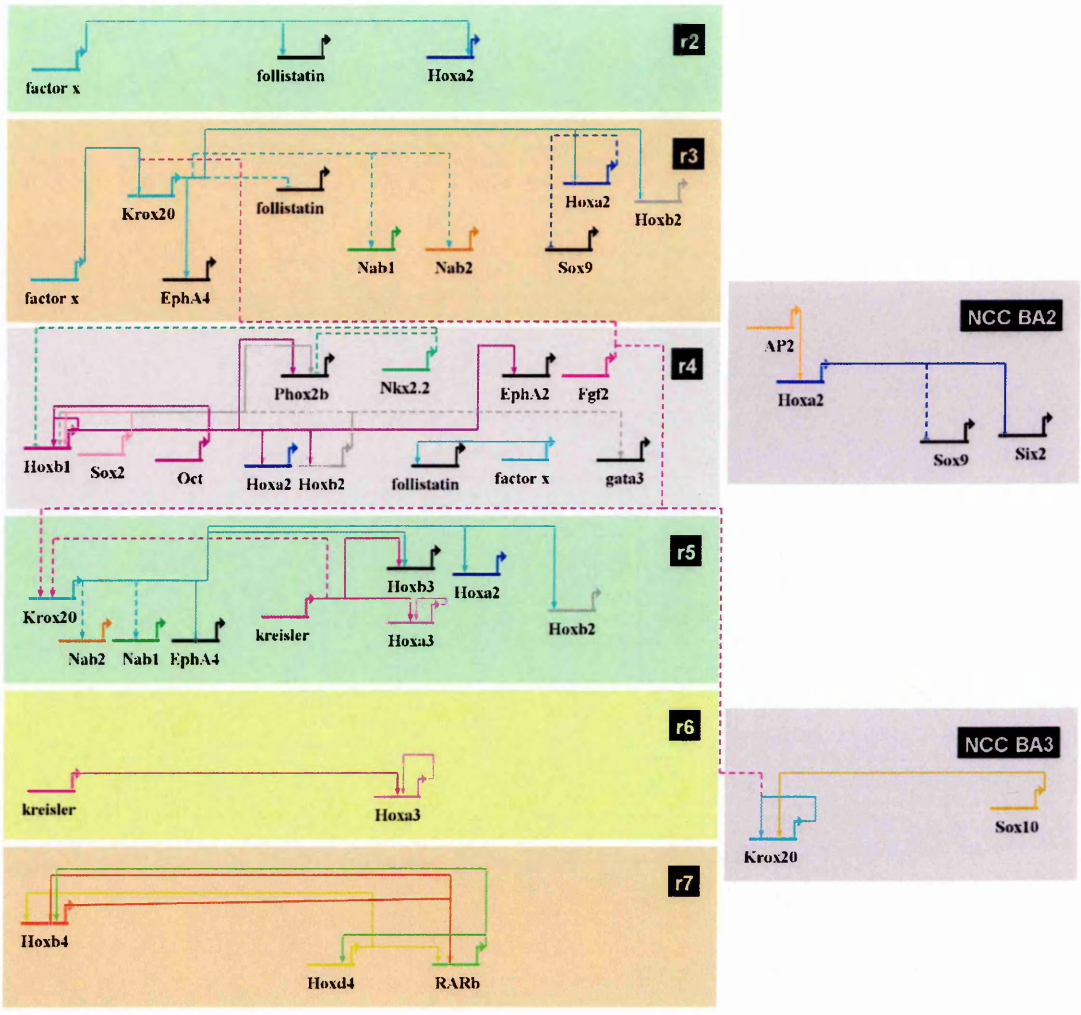


Figure 1.6. Genetic regulatory networks in the hindbrain and branchial regions at 9.5dpc.
At this stage the hindbrain can be subdivided into seven regions (rhombomeres). Dotted line illustrates that the interaction between the genes have been not been shown to be direct. Arrow head at the end of the line represents activation, whereas perpendicular line at the end of the line illustrates repression between the connected genes.

1.4.2 Intermediate Genetic Network (8 - 8.5 dpc)

During the 8 - 8.5 dpc time frame, *Hoxa1* expression regresses in the hindbrain. In contrast, *Hoxb1* expression is maintained by a number of transcription factors, and this expression becomes divided into two domains (Figures 1.4, 1.5B). The expression in pre-r4 becomes confined to r4 (Wilkinson et al., 1989), while the expression in the posterior half of the embryo, which includes the posterior neural tube, paraxial (somatic) mesoderm and gut, decreases (Frohman et al., 1990; Huang et al., 1998). The foregut expression has been shown to be dependent on a DR5 type RARE, located 3' of the *Hoxb1* coding region (Figure 1.5B, Huang et al., 1998).

In addition, *Hoxb1* expression in r4 is maintained by an auto-, para- and crossregulatory enhancer (ARE), which consists of four conserved sequence motifs (Pöpperl et al., 1995). This enhancer region consists of three Hox/Pbx (PH) bipartite binding sites (R1, R2 and R3) and one bipartite Prep/Meis (PM) site (Figure 1.5B). The R3 site makes the largest contribution to the B1-ARE regulatory activity (Marshall et al., 1992; Marshall et al., 1994; Pöpperl et al., 1995). It has been further shown that binding of a Sox2/Oct heterodimer to the ARE is required for *Hoxb1* enhancer activity in r4 (Di Rocco et al., 2001).

PH binding sites are activated by cooperative binding of Hoxb1, Hoxa1, and Hoxb2 with a member of the Pbx family (Gavalas et al., 2003; Pöpperl et al., 1995; Studer et al., 1998). Autoregulation is thus mediated by Hoxb1, and early para-regulation is mediated by Hoxa1 and late cross-regulation by Hoxb2 (Figure 1.5B).

The vertebrate *Pbx* gene is homologous to the *Drosophila extradenticle (exd)* gene. There are three members of the vertebrate Pbx family, including Pbx1, Pbx2, and Pbx3. Pbx/Exd proteins interact with a subset of *Hox* genes, in a cooperative, sequence-specific heterodimeric manner. This interaction modulates the affinity of Hox proteins for potential binding sites and activates promoters containing Pbx-responsive elements (Chang et al., 1996). The consensus binding sequence for Hox/Pbx is DGATNNATBR. The PH binding sites of the more 3' *Hox* genes *(Hoxb1* and *Hoxb2)* have a GC rich central core, whereas the PH sites of the more 5' *Hox* genes *(Hoxb4* and *Hoxb6)* have a more TA rich central core (Manzanares et al., 1999b).

The PM consensus sequence, BTGTCA, has been characterized in the regulatory regions of *Hoxb1*, *Hoxb2* and *Hoxa3* (Maconochie et al., 1997; Manzanares et al., 2001; Pöpperl et al., 1995). Meis and Prep belong to the TALE (Three-Amino acids Loop-Extension) subfamily of homeodomain proteins (Burglin, 1997). Interaction of Pbx/Exd with TALE protein family members modifies both the transcriptional activity and subcellular localization of *Hox* genes (Mann and Abu-Shaar, 1996). Prep/Pbx (PP) complexes have the ability to bind to both PH and PM motifs, allowing the formation of Prep/Pbx/Hox ternary complexes *in vitro* on bi-

partite PH motifs (Ferretti et al., 2000). The homeodomains of both Pbx and Hox are required for the interaction between these two proteins. In addition, a sequence of 20 amino acids C-terminal from the Pbx homeodomain and a conserved pentapeptide sequence N-terminal to the Hox homeodomain is essential for the interaction between Pbx and Hox proteins (Chan and Mann, 1996).

Several direct and indirect downstream targets of *Hoxb1* have been identified (Figures 1.4 and 1.6). The r4 regulatory region of *Hoxb2* has been shown to contain conserved bipartite binding sites for PH and PM (Figure 1.5D). *Trans*-activation experiments in mice confirm that these sites respond to Hoxb1 overexpression, and *in vitro* binding studies have demonstrated that Hoxb1 is able to interact with these sites, in cooperation with Pbx and Meis (Ferretti et al., 2000; Maconochie et al., 1997). Therefore, Hoxb1 regulates *Hoxb2* expression in r4, and Hoxb2, in turn, feeds back to maintain expression of *Hoxb1* in this rhombomere (Gavalas et al., 2003).

Another downstream target of Hoxb1 is the zinc-finger transcription factor *Gata3*, whose expression is downregulated in Hoxb1 null mutants (Figure 1.6, and see also Pata et al., 1999). Furthermore, the transcription factor *Phox2b* and the receptor tyrosine kinase *EphA2* are directly regulated by Hoxb1 (Chen and Ruley, 1998; Samad et al., 2004). A direct interaction between Hoxa1 and the regulatory region of *Epha2* have been also described (Chen and Ruley, 1998).

At this stage (4-6s), *Krox20* expression is observed in pre-r5. The *Krox20* expression domain in pre-r3 and pre-r5 has irregular borders, with some *Krox20* expressing cells observed in the presumptive even-numbered rhombomeres. This expression pattern sharpens at late 8.5 dpc (Irving et al., 1996). Also, NCC migrating into the third branchial arch express Krox20 (Figure 1.6). This is probably initiated by Fgfs, as it has been shown that ectopically applied Fgfs induce Krox20 expression in NCC (Marin and Charnay, 2000). It has been further shown that *Krox20* expression in the NCC is dependent on Sox10. Sox10 binding sites have been identified in the *Krox20* regulatory region, indicating that the regulatory interaction between

these two genes is direct (Figure 1.6 and see also Ghislain et al., 2003). After initiation by Fgfs, the expression of *Krox20* is maintained by an autoregulatory loop in NCC (Ghislain et al., 2003). The expression of *Krox20* in r5 is also dependent on the basic domain-leucine zipper transcription factor kreisler, since, in the *kreisler* mutant, *Krox20* expression is seen in pre-r3, but not in pre-5 (Figure 1.6 and Manzanares et al., 1999b).

The second paralogous group genes, *Hoxa2* and *Hoxb2*, are both expressed in r3 and r5. Expression of both genes in r3/5 is directly regulated by Krox20 (Figures 1.5C, D and Maconochie et al., 1999; Maconochie et al., 2001; Nonchev et al., 1996b; Sham et al., 1993; Vesque et al., 1996). Several conserved elements (RE1-RE4), which are important for *Hoxa2* rhombomeric expression, have also been identified in the *Hoxa2* r3/5 regulatory region (Figure 1.5C). Furthermore, a BoxA motif has been characterized in this region. This BoxA motif have been also identified in the Krox20-dependent regulatory region of *EphA4* and *Hoxb2* (Figures 1.5C, D, 1.6 and Theil et al., 1998; Vesque et al., 1996). Deletion of the *Hoxa2* BoxA motif results in the loss of r3 expression (Maconochie et al., 2001; Vesque et al., 1996).

Both genes, *Hoxa2* and *Hoxb2*, also show expression in r4-derived NCC migrating into the second branchial arch, but their expression is governed by different mechanisms. *Hoxb2* expression in NCC is regulated by cross-regulation of Hoxb1 and Pbx/Prep cofactors (Figure 1.5D). *Hoxa2* expression in NCC is dependent on a number of different regulatory elements (NC1-NC4), one of which has been shown to be bound by the transcription factor AP2 (Figure 1.6 and Maconochie et al., 1999; Maconochie et al., 1997). The expression of *Hoxb2* in NCC is more transient, whereas the expression of *Hoxa2* in these cells is maintained during the course of hindbrain development.

During this time, the expression of members of the third paralogous group, *Hoxa3*, *Hoxb3* and *Hoxd3*, has an anterior limit at the r4/5 border (Figures 1.3 and 1.6). The expression of *Hoxa3* and *Hoxb3* is directly initiated by kreisler (Figure 1.6 and Manzanares et al., 2001; Manzanares et al., 2002). The initiation of *Hoxb3* is also dependent on Krox20 (Figure 1.6 and Manzanares et al., 2002). *Hoxa3* is upregulated in both r5 and 6, and this expression is

maintained during hindbrain development by a conserved auto- and cross-regulatory region, dependend on the cofactors Pbx and Prep/Meis (Figures 1.5, 1.6 and Manzanares et al., 2001). In contrast, *Hoxb3* is initially expressed only in r5, but the expression is not maintained in this rhombomere.

1.4.3 Late genetic network (8.5 - 9.5 dpc)

At late 8.5 dpc (ten somites), *EphA4* is activated in r5, while its expression in r4 is downregulated (Irving et al., 1996; Nieto et al., 1992), therefore at this stage it shows strong expression in r3 and r5 and lower expression in r2 (Irving et al., 1996). At 9.0 dpc, *kreisler* expression is downregulated in both r5 and r6. In addition, it has been shown that *Hoxa2* downregulates *Sox9* expression in the second branchial arch during this period (Figure 1.6 and Kanzler et al., 1998).

The group 4 genes, *Hoxa4*, *Hoxb4*, and *Hoxd4*, have an anterior limit of expression at the r6/7 border, but the boundary of *Hoxc4* expression has a more posterior limit (Figures 1.3, 1.6 and Geada et al., 1992). All of these genes respond to RA, and a direct response has been demonstrated for *Hoxb4* and *Hoxd4* (Gould et al., 1998; Pöpperl and Featherstone, 1993). In *Hoxb4*, a 3' RARE has been identified, which is required for regulating early expression up to the r6/7 boundary, and is termed the early neural enhancer (ENE) (Figure 1.5F and Gould et al., 1998). This RARE is activated by mesodermal signals at 8.5dpc. Another autoregulatory element, which is termed the late neural enhancer (LNE), maintains this initiated expression (see very late network, Figure 1.5F and Gould et al., 1997). Furthermore, a 5' RARE has also been described in the regulation of *Hoxd4* expression (Figures 1.5G, 1.6 and Pöpperl and Featherstone, 1993).

1.4.4 Very late network (9.5 - 11.0 dpc)

At 9.5 dpc, *Krox20* expression is first downregulated in r3 and then in r5 (Irving et al., 1996). *EphA4* and the co-repressors *Nab1* and *Nab2* are also downregulated, probably as a di-

rect consequence of the downregulation of their upstream regulator, Krox20 (Figures 1.4, 1.6 and Kanzler et al., 1998; Mechta-Grigoriou et al., 2000; Theil et al., 1998).

At 9.5 dpc, the late neural enhancer (LNE) of *Hoxb4* takes over the early enhancer activity of the ENE, in order to maintain expression (Figure 1.5F). This is achieved through the action of a conserved auto- and cross-regulatory element, which consists of a single pair of Hox binding sites (TAAT/ATTA) (Gould et al., 1997). Hoxb4 itself is able to bind to this element, but it has also been shown, genetically, to be dependent on Hoxd4 (Figure 1.5F). The ENE is located adjacent to a distal of the *Hoxb3* promoter, and is shared between both genes (Figure 1.5F and Gould et al., 1998). Since *Hoxb3* is downregulated at 10.5 dpc in r5, due to the downregulation of *kreisler*, it is thereafter only expressed up the r6/7 level mediated by the shared *Hoxb4* Hox/Pbx site (Gould et al., 1997).

1.5 Loss-of-Function and Gain-of-Function analysis of *Hox* genes and their regulators

The development of gene targeting ('knock out') technology has allowed for the deletion or replacement of endogenous genes (Capecchi, 1989). This technology is based on 'homologous recombination in mouse embryonic stem (ES) cells, which are then introduced into the mouse germ line by blastocyst injection.

1.5.1 Mutants of the RA pathway

1.5.1.1 *RAR* and *RXR*

The RA receptors have been inactivated in mouse gene targeting experiments. In general homozygote mutant embryos for each of the receptors do not show any overt phenotype and have normal hindbrains (Mendelsohn et al., 1994b). Whereas, compound mutants of *RARα/RARβ, RARα/RARγ and RARγ/RARβ* show severe phenotypes. In the *RARα/RARβ* double mutants, r6 and r7 are fused (due to the loss of the r6-r7 boundary) and show an expanded r5 (Dupé et al., 1999). In *RARα/RARγ* compound mutants, no rhombomere boundaries formed

(Wendling et al., 2001). Also, in this double mutant the caudal hindbrain has acquired an anterior character, as it expresses a combination of r3 and r4 molecular markers such as Krox20 and Hoxb1, instead of expressing kreisler. These observations indicate that RARα and RARγ mediate the RA signaling pathway for establishing the identity of r5 and r6.

In addition to the *RAR* mutants, *RXR* mutants have been generated and analyzed as single as well as compound mutants (Kastner et al., 1997a; Kastner et al., 1997b; Krezel et al., 1996; Sucov et al., 1994; Sucov et al., 1995). Single mutants for *RXRγ* and *RXRβ* are phenotypically normal, whereas *RXRα* null mutants are embryonic lethal between 13.5 dpc and 16.5 dpc when bred to homozygosity, as a result of heart defects (Sucov et al., 1994). This phenotype is also observed as an effect of embryonic vitamin A deficiency.

By generating different combination of *RXR* compound mutants it was shown that RXRγ can be fully substituted by functional RXRα and RXRβ. Further it was demonstrated that RXRα has the most important function of the RXRs and can compensate for the other *RXR* null mutants (Krezel et al., 1996). This implies that there is a large degree of redundancy amongst RXRs. This situation is different from that of RARs, since all types of *RAR* compound mutants show much broader effects than single mutants (Dupé et al., 1999; Mendelsohn et al., 1994b).

The fact, that mice lacking a single *RAR* or *RXR* gene have only a limited number of defects, can be explained that upstream targets of RA just require a limited level of RAR and RXR heterodimer formation. This is confirmed by analysis of different combinations of *RAR/RXR* compound mutants (Kastner et al., 1997a), which recapitulate the defects seen in *RAR* double mutants. This data implies that RXR (especially RXRα) and RAR interaction is essential for normal embryonic development.

1.5.1.2 RALDH2

Homozygous mutant embryos for *RALDH2* die at approximately 10.5 dpc and the segmentation, growth and patterning of the hindbrain is affected (Niederreither et al., 1999;

Niederreither et al., 2000). The transcription factor *Krox20*, normally expressed in r3 and r5, shows expression in a single domain in the whole caudal hindbrain of mutant embryos. *Hoxb1* which is upregulated in wildtype embryos in r4 is expressed in the caudal most of the hindbrain of the mutant embryos. Furthermore, *kreisler*, which normally defines r5 and r6, is absent in *RALDH2* mutant embryos. Altogether, the posterior hindbrain appears to be anteriorized, as more anterior expression markers are expressed in more posterior domain and posterior markers are absent or downregulated.

1.5.2 Mutational analysis of *Hox* genes regulators

1.5.2.1 Krox20

Null mutants for *Krox20* die at birth and exhibit several defects in the hindbrain (Schneider-Maunoury et al., 1993; Swiatek and Gridley, 1993). The presumptive regions of r3 and r5 are formed during early development, but they are not maintained later in development. The total length of the rhombencephalon is reduced. This has an effect on the trigeminal ganglion, which is fused with facial and vestibular ganglia. The abducens (VIth) nerve, derived from r5, is absent as a result of the reduction in the size of r5.

At the molecular level, it has been shown that, in *Krox20* deficient mice, *Hoxa2*, *Hoxb2, Hoxb3*, and *EphA4* are downregulated (Seitanidou et al., 1997). This observation is confirmed by previous studies, which show that Krox20 directly regulates these genes (Manzanares et al., 2002; Nonchev et al., 1996b; Sham et al., 1993; Theil et al., 1998). Further studies of these mutant mice have revealed that Krox20 is involved in repression of the *follistatin* gene in r3, but not in r5. These studies show that Krox20 is an important regulator in patterning the hindbrain through its activation and repression of other key molecular players.

It has been shown in compound mutant of *Krox20* and *Hoxa1*, that both genes synergize in a dosage-dependent manner in r3 patterning (Helmbacher et al., 1998 2295). This is rather surprising, since r3 is not within the *Hoxa1* expression domain. This suggests that

Hoxa1 indirectly regulates *Krox20*, and the development of r3 is governed in a non-autonomous way.

1.5.2.2 kreisler

The *kreisler* mutant was identified in an x-ray induced mutagenesis screen (Deol, 1964). The mutant mice show inner ear, vestibular defects and defects in neural crest-derived skeletal elements (Deol, 1964; Frohman et al., 1993). In addition, the hindbrain is affected, particularly in the otic region (Manzanares et al., 1999a; McKay et al., 1994). This explains the circling behavior, seen in the mutant mouse, which is characteristic of inner ear and vestibular defects. The rhombomeric borders between r4, r5, and r6 are not present, leading to a thickening of the posterior hindbrain neuroepithelium. r5 is absent, but r6 has not lost its character (Manzanares et al., 1999a). This is confirmed by the absence of r5 specific genes, including *Krox20, Hoxb2, Hoxb3* and *Hoxb4*. The expression of *Fgf3*, *CRABP*, and *Hoxa3* in r6 is absent, which suggest that r6 has not fully developed. This also implies that kreisler has an important role to play in regulating segmental identity in r5. This view is supported by ectopic expression of kreisler in the hindbrain, which leads to a change of an r3-like identity into an r5-like identity (Theil et al., 2002).

Loss of r5 also has consequences on the fate of the cranial nerve. The abducens (VI) nerve is not present, and the additional absence of the glossopharyngeal (IX) ganglion and nerve suggest that r6 has partially lost its identity. This is further supported by the fact that other neural crest-derived cartilaginous structures, derived from r6, are not affected in the mutant (Cordes and Barsh, 1994; Frohman et al., 1993).

In addition to the hindbrain malformations, the hyoid bone in the adult neck region is affected. The greater horn contains additional appendages, which resemble the lesser horn of the hyoid bone. These data can be explained by NCC migration. A broad stream of NCC, from r4 up to r7, migrates in both the second (lesser horn of the hyoid) and third (greater horn of the hyoid) BAs. The hyoid phenotype probably results from the migration of r6 derived NCC into

the second BA, due to the intermixing of r4- and r6-derived NCC (Frohman et al., 1993; Manzanares et al., 1999a).

1.5.3 Hox mutants

To further understand the functions of the *Hox* genes, a number of these genes have been either deleted from their endogenous locus or gain-of-function mutants have been generated in mice. In general, the gene disruption data suggests that regional identity along the A-P axis is governed by the most posteriorly expressed *Hox* gene in that region. This idea has lead to the proposal of the "posterior prevalence" model (Chisaka and Capecchi, 1991; Chisaka et al., 1992; Duboule and Morata, 1994; Lufkin et al., 1991), which states that more posterior *Hox* genes in the cluster tend to be more prevalent than the more anterior genes. This idea is further supported by ectopic overexpression of posteriorly expressed *Hox* genes in more anterior regions, which leads to posterisation of regions (Bell et al., 1999; Cho et al., 1991; Kessel and Gruss, 1990; Zhang et al., 1994).

In summary, posterior transformations are observed in gain-of-function experiments, whereas, in loss-of-function mutations, anterior transformations occur. However, not all *Hox* mutants follow this rule (Jegalian and De Robertis, 1992; Pollock et al., 1992), and further analysis must be performed to understand the molecular basis of this model. Almost all *Hox* genes have been disrupted and a variety of embryonic tissues are affected, including neural ectoderm, neural crest, paraxial mesoderm, placodal ectoderm and limbs.

1.5.3.1 *Hoxb1* and *Hoxa1*

Although the generation of rhombomeres in Hoxb1 mutants is not affected, their respective identity is altered (Goddard et al., 1996; Studer et al., 1996). The expression of *Krox20* or *kreisler* appears normal in the absence of *Hoxb1*, but molecular markers including *Wnt8* and *Hoxb3*, are not upregulated in r4 at later stages. Additional work has shown that two subpopulations of neurons (facial branchiamotor neurons and contralateral vestibuloacoustic

efferent neurons) are unable to migrate into the correct location, and, therefore, there is a loss of the facial motor nerve.

By testing the r2 specific enhancer (from the regulatory region of *Hoxa2*) in the *Hoxb1* mutant, it was possible to show that the r4 identity is changed to an r2 identity. This is further confirmed in *Hoxb1* mutants, in which the migratory behavior of the r4 neurons adopts a more lateral movement similar to the trigeminal neurons from r2. Hoxb1 also plays an important role in the development and maintenance of the VIIth cranial nerve; this was shown by conditionally eliminating *Hoxb1* expression in r4 derived NCC (Arenkiel et al., 2004).

A gain-of-function study using retroviral misexpression combined with orthotopic grafting in chick embryos supports this idea (Bell et al., 1999). When *Hoxb1* is exclusively overexpressed in r2, the r2 motor axon projections are reorganized to resemble r4 motor axons (Bell et al., 1999). Altogether, Hoxb1 plays an important role in maintaining r4 identity and in controlling the migration of the motor neurons in the hindbrain.

Hoxa1 mutants die at birth due to anoxia. In these mutants the inner ear and cranial nerves are affected (Lufkin et al., 1991). In addition, hindbrain development is severely disturbed, r4 is reduced in size, r5 is missing; and defects in r5 up to 8 are apparent. Furthermore, cells with an r2-like identity are present in r3, which probably explains why the r3 motor nerves show migration behavior similar to even-numbered rhombomeres. This data suggest that unlike Hoxb1, Hoxa1 functions both as a segmentation and identity gene.

In compound mutations of *Hoxa1* and *Hoxb1*, it has been also shown that Hoxb1 has in role in the normal development of tissues derived from r4 neural crest (Gavalas et al., 1998; Rossel and Capecchi, 1999). In this compound mutant the r4-derived neural NCC are not able to develop or migrate correctly, as a consequence all second-arch derivatives are missing (Gavalas et al., 2001). Thus the double mutants, shows a variation of phenotypes in the hindbrain which are not present in either of the single mutants (Gavalas et al., 1998; Studer et al., 1998), implying that there is a synergistic effect between Hoxa1 and Hoxb1.

1.5.3.2 *Hoxa2* and *Hoxb2*

Hoxa2 homozygous mutant mice died at birth, while mice heterozygous for the *Hoxa2* mutation appeared normal. In *Hoxa2* null mutants, patterning of the BA is affected (Gendron-Maguire et al., 1993; Rijli et al., 1993). The more anterior hindbrain regions are also affected, which is expected since *Hoxa2* is the only *Hox* gene expressed in r2 (Krumlauf, 1993; Prince and Lumsden, 1994).

In the null mutant embryos, the second pharyngeal (hyoid) arch NCC-derived cartilaginous elements are absent (Figure 1.7B). They lack the lesser horn of the hyoid bone, the stapes, and the styloid bones. Instead they have abnormal skeletal pieces, membranous and endochondral in origin, which are duplications of proximal bones normally present in the first mandibular arch. Therefore, in the absence of *Hoxa2* activity, second arch neural crest cell fate is transformed into a first arch identity. The selector gene property of *Hoxa2* has also been confirmed by overexpression experiments in chicken and *Xenopus* (Grammatopoulos et al., 2000; Pasqualetti et al., 2000). In addition to skeletal defects observed in the *Hoxa2* mutant, other BA structures are affected. The ear pinna, which derives from the first and second arches, is absent and replaced by an ill-defined protuberance. Also, the trajectories and insertions of the muscles of tongue and hyoid are abnormal (Gendron-Maguire et al., 1993).

Overexpression of *Hoxa2* in *Xenopus* in the first arch suppresses jaw formations and transforms the first arch neural crest derivatives into structures typical of the second arch (Pasqualetti et al., 2000). Similar results have also been observed in chick overexpression experiments (Grammatopoulos et al., 2000). This is the opposite phenotype of the murine *Hoxa2* mutant. Therefore, *Hoxa2* acts as a selector gene in the second BA.

The function of Hoxa2 has been further investigated in the hindbrain region at both the molecular and morphological level in *Hoxa2* deficient mice, where it plays an important role in controlling r2-r3 motor axon pathfinding and establishing rhombomeric boundaries (Figure 1.7A and Gavalas et al., 1997). Since r2 is transformed into an r1 like identity, the cells fated to contribute the alar regions of r2 and r3 are reduced, and r1 is extended. In addition, the r1/2 boundary is absent. The ectopic expression of specific markers, such as *Sax-1* and *En-2* and the absence of *EphA4* in r2 suggest a homeotic transformation of r2 into r1 (Gavalas et al., 1997). Furthermore the axonal projections of the trigeminal (V) and facial (VII) motor nuclei are abnormal in r2 of the *Hoxa2* deficient mice. Further detailed analysis shows that the mutant cerebellum is extended, and the cochlear nuclei column is reduced. This is consistent with the observation that r1 and the isthmic region generate the majority of the cells of the cerebellum, whereas r2 has a limited contribution.

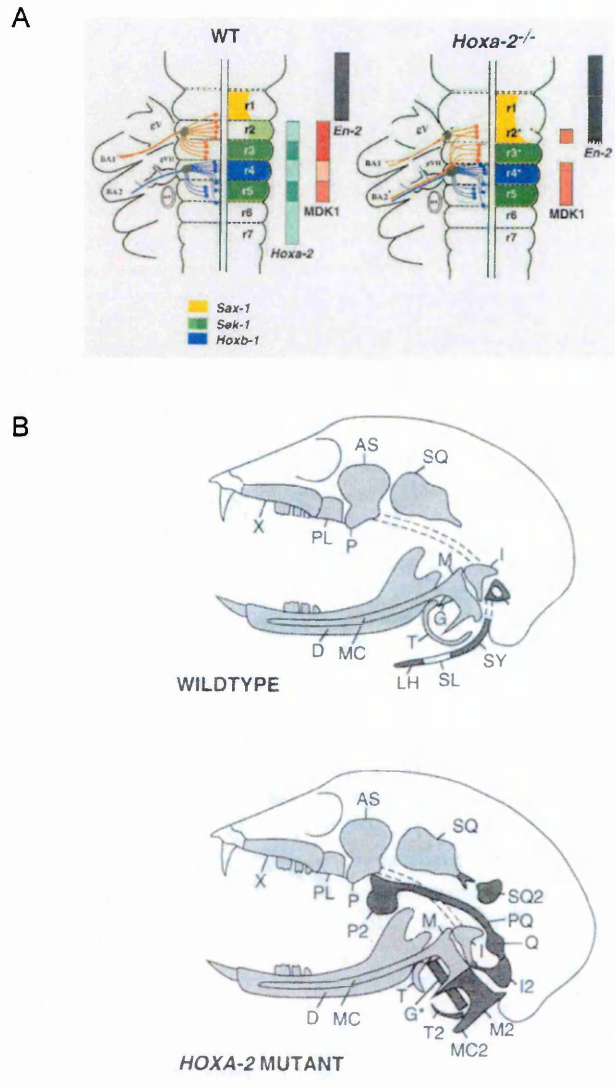


Figure 1.7. Comparisons of the craniofacial structures between a wildtype and a *Hoxa2* mutant.
(A) Schematic diagram of the expression domains of various genes in the hindbrain of a wildtype embryo and of the *Hoxa2* mutant (Gavalas et al., 1997). **(B)** Schematic diagram of the skeletal elements derived from BA1 (gray) and BA2 (dark) in wildtype and *Hoxa2* mutant mice. In the wildtype, the stapes (S), styloid bone (SY), stylohyoid ligament (SL) and the lesser horn of the hyoid bone (LH) are derivatives from the cartilage of the second arch (Reicherts cartilage). The *Hoxa2* mutant exhibit duplications of structures normally present in the first mandibular arch: an extra incus, tympanic, a supernumerary malleus fused to an ectopic truncated Meckel's cartilage, and squamosal bones (Rijli et al., 1993).

Hoxa2, therefore, not only acts as an selector gene for second arch mesenchymal neural crest cells, but it also plays an important role in patterning the anterior hindbrain region by establishing the identity of r2 and, to a lesser extent r3. A hypomorphic *Hoxa2* mutant, which generates lower levels of Hoxa2 protein, shows that the second BA is rapidly affected by low level of transcripts, whereas the hindbrain remains normal even at low levels of *Hoxa2* transcripts (Ohnemus et al., 2001).

Different *Hoxb2* null mutants have been generated, which exhibit variation in their phenotypes (Barrow and Capecchi, 1996; Davenne et al., 1999). One mutant strain showed a wide effect on the expression of other genes in *cis* (Barrow and Capecchi, 1996). For example, *Hoxb1* expression is eliminated in r4. On the other hand, Davenne et al., described a *Hoxb2* mutant and different effects have been observed (Davenne et al., 1999). In this case, the single *Hoxb2* null mutant the identity of r4 was changed in later stages, supporting the idea that it is maintaining *Hoxb1* expression (Gavalas et al., 2003) .

In addition to these A-P changes, also dorsoventral (D-V) changes have been observed. Different markers, including *Math3, Nkx2.2* and *Phox2b* are upregulated in r4, which resembles normal expression in r2 (Davenne et al., 1999). Therefore, *Hoxb2* has an important role for motorneuron development in the ventral basal plate of r4. This has been confirmed by retrograde dye tracing of r4 branchiomotor neurons, which were reduced and projecting incorrectly to the r4 exit point (Davenne et al., 1999).

In the *Hoxa2* and *Hoxb2* compound mutant inter-rhombomeric boundaries are missing. Furthermore, the expression of *Pax3* and *Pax6*, which are expressed in D-V domains and are important for differential determination of neural progenitors, is altered in r2 and r3. It has been also shown in this compound mutant that certain ventral interneuron subtypes are missing in r3 (Davenne et al., 1999).

1.5.3.3 Hoxa3, Hoxb3 and *Hoxd3*

The *Hoxa3* null mutant shows severe organ defects, including failure of the thymus, parathyroid and thyroid tissue glands development (Chisaka and Capecchi, 1991). Also, the vessels and musculature of the heart and the most posterior bones in the face are affected. All of these structures are derivatives of mesenchymal NCC, however derivatives of neurogenic NCC are also affected. The mutant phenotype can be distinguished into two classes according their penetrance. The first class of mutant phenotype involves a deletion of the proximal portion of the IX (glossopharyngeal) cranial ganglion which fails to connect to the hindbrain (Manley and Capecchi, 1997). In the second class of mutant phenotype a fusion of the IXth and IXth cranial ganglion was observed.

The *Hoxb3* null mutants has at minor defects in the formation of both the cervical vertebrae and the IX cranial nerve (Manley and Capecchi, 1997). As with the *Hoxa3* null mutant both classes of phenotype are visible but at a lower penetrance.

Targeted disruption of *Hoxd3* leads to the homeotic transformation of the second cervical vertebrae (the axis) to the first cervical vertebrae (the atlas) (Condie and Capecchi, 1993). The anterior arch of the atlas is also transformed to an extension of the basioccipital bone of the skull. The affected structures are all derived from paraxial and lateral mesoderm. However, there is no phenotype in the cranial nerves (Condie and Capecchi, 1993; Manley and Capecchi, 1997). On the other hand, the generation of compound mutants *(Hoxa3/Hoxb3, Hoxa3/Hoxd3* and *Hoxb3/Hoxd3)* shows that *Hoxd3* has a synergistic role in ganglia formation with *Hoxb3*, since there is an increase in the penetrance of the IXth cranial ganglion defects in *Hoxb3/d3* double mutants (Manley and Capecchi, 1997). Similar synergistic interactions are deduced from the *Hoxa3/Hoxd3* compound mutant (Condie and Capecchi, 1994). This double mutant reveals a more dramatic phenotype as supposed by the additive defect resulted from both mutants. The entire atlas is missing rather then undergoing an extensive homeotic transformation into the more anterior vertebrae. This result suggests that both genes are interacting to differentially regulate the proliferation rates of the affected structures and further that that these

paralogous genes operate in multiple tissues in the same region, but also that they appear to be performing the same functions in these tissues.

1.5.3.4 *Hoxa4, Hoxb4, Hoxc4* and *Hoxc4*

Hoxa4 mutants show a partial transformation of the third cervical vertebra (C3) to the second cervical vertebra (C2) as well as a cervical rib at C7 (Horan et al., 1994). *Hoxb4* null mutants display partial homeotic transformations of C2 to C1, defects in the closure of the sternal rudiments, and variably penetrant neonatal lethality (Ramirez-Solis et al., 1993). *Hoxd4* homozygous or heterozygous deficient mice have more complete homeotic transformations of C2 to C1 as those seen in Hoxb4 mutants, and also malformations of the neural arches of C1 and ectopic ossification at the basioccipital bone (Horan et al., 1995). Also the phenotypes of compound mutants with different combinations of these genes have been examined (Horan et al., 1995). In the compound mutant *Hoxb4/Hoxd4* the C2 to C1 transformation become completely penetrant. In the triple mutant *Hoxa4/Hoxb4/Hoxd4* additional vertebrae are transformed, with C5 partially adopting the identity of C1. This implies that these genes are required in establishing the identity of a larger anterior-posterior domain. In all of these mutants the anterior transformation were at the respective anterior borders of expression of each of these genes, suggesting that the genes of the fourth paralogous group have a crucial role in establishing the identity of the second and third cervical vertebrae.

1.6 Aims of my projects

Several studies have shown that *Hox* genes have a common role in embryonic development in multicellular organism. This observation has been confirmed by their genomic organization, expressions pattern and mutational analysis. Further it has been shown that the mechanism of *Hox* genes regulation is highly conserved. This allows me to use different model systems and, also, to compare genomic information from different species to understand the regulatory mechanism of *Hox* gene regulation. One approach of understanding the regulation of genes is to analyze the *cis*-regulatory elements that regulate the spatial-temporal *Hox*

expression and identify *trans*-acting factors (transcription factors) which mediate this expression. Here I am using chick and mouse embryos as model systems and focus on the regulation of *Hox* gene expression in the hindbrain. The aim of this project is to understand the regulatory mechanism for *Hoxa2* and *Hoxb1* expression in the hindbrain.

Hoxa2 expression in the hindbrain is upregulated up to the r1/2 border and in NCC migrating into the second BA. It is the only *Hox* gene expressed in r2. Mutational analysis has shown that it plays a crucial role during cranialfacial development. However, the regulatory mechanisms of *Hoxa2* are only understood in r3/5 and NCC, and remain elusive in r2 and r4. Here I use chick electroporation and transgenic mouse analysis and genomic sequence comparisons to investigate the molecular regulatory mechanism required for the restricted temporal and spatial domains of *Hoxa2* in rhombomeres 2 and 4. In Chapter 3, I will concentrate on the regulatory mechanism directing r4 expression of *Hoxa2* in the hindbrain. Chapter 4 of this report will concentrate on the molecular mechanism required for r2 expression. Chapter 5 describes the mechanism which explains the differential expression of the duplicated *Hoxa2* genes, *Hoxa2(a)* and *Hoxa2(b)* in fugu.

Hoxb1 has been shown to be important for establishing r4 identity. Its expression is restricted to r4 during development. My aim is to understand the molecular mechanism for r4 restriction. In Chapter 6, I focus on a region which has r3/5 repressor activity, therefore eliminating *Hoxb1* expression in r3 and 5. In Chapter 7, I investigate the regulatory region which mediates the repression of *Hoxb1* in the second BA. Finally, in Chapter 8 I will present data which arose by collaborative work.

Chapter 2

Materials and Methods

Molecular biological techniques were based on those described by Sambrook et al. (1989) and the suppliers' instructions. Oligonucleotides were supplied by Invitrogen and Integrated DNA Technologies (IDT). Enzymes were obtained from New England Biolabs (NEB), Promega, Boehringer Mannheim, and Invitrogen, unless otherwise stated. Buffer solutions and media were supplied by the Core Facility of SIMR. All the oligonucleotide sequences are listed in 5' to 3' orientation.

2.1 Standard solutions and reagents

Ampicillin

Dissolved at 50mg/ml in sterile water, filter sterilized, and stored at -20°C. Added to medium and agar to a final concentration of 50μg/ml.

Avertine

10g 2,2,2-Tribromoethanol dissolved in 10ml tertiary amyl alcohol. Stored in the dark at 4°C.

Fixative

1% Formaldehyde, 0.2% Gluteraldehyde, 2mM MgCl2, 5mM EGTA, 0.02% Nonidet P40 (NP40)

Wash solution

0.02% NP40 in PBS

Stock Substrate

40mg/ml 5-Bromo-4-chloro-3-indodyl-β-D-galactopyranoside (X-gal) dissolved in dimenthylformamide stored at -20°C.

Staining solution

5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$-3H2O,2mM $MgCl_2$, 0.01% Sodium deoxycholate, 0.02% NP40, 1mg/ml from stock X-Gal. Stored in the dark at 4°C.

L-Broth (LB)

1%(w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, and 0.5% (w/v) NaCl. Sterilized by autoclaving and stored at 4°C.

L-agar

LB containing 1.5% (w/v) bacto-agar. Autoclaved, cooled to 55°C, added antibiotic for selection, poured to plates and dry at room temperature. Stored at 4°C.

TE (Tris-EDTA) buffer

10mM Tris-HCl (pH 8.0), 1mM EDTA (pH 8.0). Sterilized by autoclaving

Injection buffer

10mM Tris-HCl (pH 7.6), 0.1mM EDTA

NZCYM

10g NZ amine, 5g Bacto-yeast extract, 5g NaCl, 1g casamino acids and 2g of MgSO4-$7H_2O$ were dissolved in 900ml of water. The ingredients were stirred until everthing had dissolved. The pH was adjusted to pH 7.0 with NaOH. Following this, the volume was adjusted to one liter and sterilized by autoclaving.

PBS (x10)

1.3M NaCl, 70mM Na2HPO4, 30mM NaH_2PO_4 , pH adjusted to 7.0 and autoclaved.

SSC (X20)

3M NaCl, 0.3M sodium citrate pH adjusted to 7.0

Hybridization buffer (for Phage screening)

12,5ml 20xSSC, 5ml 10x BLOCKING SOLUTION (Roche), 0.5ml 10%N-Lauryl, 0.1ml 10% SDS, 31.9ml water

Hybridization buffer (for Southern Gel transfer)

0.5M $NaPO_4$ (pH 7.2), 7% SDS, 10mM EDTA

Depurination solution

0.2N HCl

Denaturation solution

1.5M NaCl, 0.5M NaOH

Neutralization solution

3M NaCl, 0.5M Tris HCL (pH 7.5)

Digestion buffer

100mM Tris HCl (pH8), 200mM NaCL, 5mM EDTA, 0.2% SDS, 0.2mg/ml proteinase K

SOC medium

0.5% Yeast extract, 2.0% tryptone, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 20mM $MgSO_4$, 20mM glucose

TBE (Tris-borate EDTA) buffer

89mM Tris base, 89mM boric acid, 2mM EDTA (pH 8.0). Made as a 10x stock.

Fast Green

Stock solution 0.5 mg/ml Fast Green (EM Science) in water. 1/20 volume fast green of the DNA solution was then added to the DNA solution and mixed by tapping and pipetting.

DEPC

Diethylpyrocarbonate (DEPC) (BDH) was added to a final concentration of 0.1% and the solution incubated overnight at room temperature. The solution was then heated in a boiling water bath for 10 minutes and sterilized by autoclaving.

2.2 DNA manipulations

2.2.1 Transformation of competent cells

DNA was transformed into cells of the *Escherichia coli* strain DH5α. 50µl of competent cells were incubated in a pre-chilled 15ml polypropylene tube (Falcon 2059), and 10µl of the ligation reaction was added. The cells were incubated on ice for 30 minutes, followed by a 2-minute heat shock at 42°C. The tube was then placed back on ice for another 2 minutes. Next 500ml of SOC medium was added, and the cells were incubated for 30 minutes with

shaking at 37°C. The transformation mixture was then spread onto L-agar plates containing the appropriate selection antibiotic and incubated overnight in an incubator at 37°C. The next day, colonies were picked and incubated overnight in a shaking incubator at 37°C for further analysis.

2.2.2 Plasmid DNA isolation

Plasmid DNA was isolated using the Fast Plasmid Mini kit (Eppendorf) according to the manufacturer's instructions. Briefly, an overnight culture (2ml) was centrifuged, the pellet was resuspended in RNase/Lysozyme lysis solution, and the cleared lysate applied to a spin membrane column. The membrane was washed to remove salts and proteins, and DNA was eluted with water. The DNA was then used for restriction analysis or sequencing. Samples were stored at -20°C.

For large-scale plasmid purification, DNA was isolated from 100ml overnight bacterial cultures using the Qiagen Plasmid Maxi Kit (Qiagen) according to manufactures instructions. Briefly, the bacteria were subjected to alkaline lysis and bound to an Anion-Exchange resin under appropriate low-salt and pH conditions. RNA, proteins, dyes, and impurities are removed by multiple washes. Plasmid DNA is eluted in a high-salt buffer and the concentrated and desalted by isopropanol precipitation. Samples were stored at -20°C.

2.2.3 Polymerase Chain Reaction (PCR)

Regions were amplified by using the PCR Master Mix (Promega), unless otherwise stated. PCR primers were diluted in water to a final concentration of 100µM. PCR reactions were set up in a total volume of 25µl.

A typical PCR reaction mixture was:

PCR Master Mix	12.5µl
primer (forward)	0.1µl
primer (reverse)	0.1µl
template	xµl

water	xμl
total volume	25μl

The DNA Engine (MJ Research) was used for amplification with different programs depending on the size of the expected PCR product and the T_M value of the primers.

2.2.4 DNA Gel extraction

PCR products and specific restriction fragments for subcloning were isolated from agarose gels using the MinElute gel extraction kit (Qiagen) according to the manufacturer's instruction. Briefly, after the DNA was run in a gel, the desired DNA band or fragment was excised from the gel using a razor blade. The gel slice was solubilized at 50°C in the buffer provided, and DNA was collected by absorption to a membrane in a centrifugation column. After washing to remove salts, DNA was eluted from the membrane in water.

2.2.5 DNA restriction enzyme digest

Most reactions were carried out in a volume of 100μl, using enzymes purchased from NEB and Promega and using the manufacturers' recommended 10x transcription buffer appropriate to the enzyme being used.

A typical reaction mixture was:

DNA	xμl
10x restriction buffer	10μl
BSA (10mg/ml	1μl
DEPC water	xμl
Restriction enzyme	3μl
total volume	100μl

Most reactions were incubated at 37°C for 2-3 hours.

2.2.6 Ligation of DNA fragments

100ng of vector DNA was ligated with a 20-50 fold molar excess of insert. 1µl of 10x Ligase Buffer (Invitrogen) was added to the vector followed by the insert DNA. 1µl of T4 DNA Ligase (1units/µl) (Invitrogen) was added along with H_2O, to bring the total volume to 10µl. The ligation reaction was incubated at 4°C overnight.

2.2.7 Cloning of annealed Oligonucleotides

For the multimerization constructs oligonucleotides were resuspended in TE (10mM Tris and 1mM EDTA) to the final concentration of 100µM. From each of both stocks 5ul were taken and diluted into 90µl of TE. This was then held into boiling water for 5 minutes and cooled down to room temperature. Then the annealed oligonucleotides were used for ligation with the appropriate vector.

2.2.8 Site-directed Mutagenesis

For mutagenesis the QuikChange Site Directed Mutagenesis Kit (Stratagene, #200518) was used. The reactions were performed in the original vector, and the insert was then cloned into the reporter expression vector (BGZ40). It was performed according to the manufacturer's instruction. Briefly, the sample reaction was set up with recommended 10x transcription buffer, template, forward and reverses oligonucleotide, dNTP mix, QuikSolution, PfuUltra HF DNA polymerase along with water, to bring the total volume to 50µl. The reaction was cycled using the cycling parameters recommended by the manufactures. The restriction enzyme *DpnI* was added to the reaction and incubated at 37°C for 1hour to digest the parental DNA. The DNA was transformed and the insert was sequenced to verify that selected clones contain the desired mutation(s).

2.3 Phage Screening

The bat *(Carollia perspicillata) Hoxa2* regulatory region was isolated from a phage library. The partially *Sau3AI* digested 15-22kb genomic fragments were cloned into *BamHI* di-

gested lambda DASHII and packaged using Gigapack III Gold. The library was simultaneously amplified and fractionated by plating 20,000pfu onto 95 plates. The library gave about five-fold genomic coverage. Phage DNA was prepared from each of 94 sublibraries and arrayed onto microtiter plates to facilitate a primary PCR-based screen. The library was provided courtesy of Chris Cretekos and Richard Behringer.

Screening primers were designed in highly conserved regions identified by alignment of the *Hoxa2* downstream intergenic regions from mouse and human. Seven primer pairs were tested on genomic bat DNA. Three primer pairs gave the expected product size and were used for screening the sublibraries. The primers had the following sequences:

1. GAACAAAYTTTCTAGTCGAG and TTC TGA ACM ATC AGG GTA
2. CCCCAATGAGGCGTTCCTT and CGA CGG CTG GCT ATT GAT
3. TTTCAACAAGTACCTTTGC and GGT GCT GAA AAT GTT TCA

The PCR products were TA-cloned (pGEM-T, Promega) and sequenced. The 96 phage sublibraries were then screened with these three primer pairs. Sublibraries, which were positive for all three PCR reactions, were used for further screening. In total, 9 of the 96 libraries were positive with all three primer pairs.

In order to isolate individual clones from the sublibraries, the positive sublibraries were grown on plates, transferred to filters, and screened with a DIG-labeled probe. For screening the phage libraries, the three PCR fragments were labeled with DIG-11-UTP (Roche, PCR DIG Probe Synthesis Kit). The PCR products were denatured in Hybridization buffer for 5 minutes at 95°C and incubated immediately on ice.

2.3.1 Preparing the Plating Cultures

50ml of LB broth, supplemented with 0.2% maltose and 10mM $MgSO_4$ was incubated in a sterile flask with a single colony of the *E. coli* strain ER1647 (NEB). The culture was grown with shaking at 30°C overnight. The next day, the cells were centrifuged in a sterile

conical tube for 10 minutes at 2000rpm. The medium was carefully decanted off the cell pellet and gently resuspended in 15ml of 10mM $MgSO_4$.

The cells were diluted to OD_{600}=0.5 with 10mM $MgSO_4$. About 600µl of cells were needed for each 150mm plate. The cells were combined with the phage library containing $5x10^4$pfu of bacteriophage. This mixture was incubated for 15 minutes at 37°C.

6.5ml of top agar, melted and cooled to 48°C, was mixed with each aliquot of infected bacteria and spread evenly onto a freshly poured bottom agar plate. The plate was incubated at 37°C for 6-8 hours and placed at 4°C for 2 hours to chill.

2.3.2 Performing Plaque Lifts and Hybridization

The plaques were transferred onto a nylon membrane, 30 sec for the primary lift and 4 minutes for the secondary lift. The position and orientation of the membrane on the plate was recorded with a syringe needle dipped in ink by stabbing through each membrane and plate agar. The nitrocellulose membranes were denatured after lifting by submerging the membranes in a 1.5M NaCl and 0.5M NaOH denaturation solution for 2 minutes, neutralized for 5 minutes by submerging in a 1.5M NaCl and 0.5M Tris-HCL (pH 8.0) neutralization solution, and rinsed for 30 seconds by submerging in a 0.2M Tris-HCL (pH 7.5) and 2xSSC buffer solution. The membranes were then blotted on 3MM paper (Whatman) and the DNA was crosslinked to the membrane using the autocrosslink setting at 1200000µJ/cm^2 of UV energy. The plates were stored at 4°C to use for phage clone recovery.

After 2 hours incubation in blocking solution at 56°C, the membranes were incubated with 4ml hybridization buffer, including Dig labeled probe, overnight. After hybridization, the membranes were rinsed twice for 5 minutes in 2xSSC, 0.1% SSC at 15-25°C and twice for 15 minutes in 0.5xSSC, 0.1 SDS (prewarmed) at 68°C in an incubator under constant agitation. The immunological detection was performed using the DIG High Prime DNA labeling and Detection Starter KitII (Roche, #1585 614). The signals on the exposed films were compared between the primary and secondary lifts. If the signals aligned, the plaque was isolated by a

pipette tip. The positive plaques were re-plated and screened again. At the end, four different phage clones were isolated.

2.3.3 Phage Mini-Prep

100µl of stationary phase bacteria were added to 6ml of NZCYM and cultured at 37°C for 30 minutes to 1 hour. The OD_{600} should be 0.2. Positive plaques were added and incubated at 37°C for 20 minutes. Following this the mixture was incubated in a shaker for 6-8 hours at 37°C at 200rpm. Next, 1 drop of $CHCl_3$ was added in order to remove bacteria. The mixture was centrifuged at 5000rpm for 5 minutes, and the upper layer was transferred.

A DNase(10mg/ml, 2µl)/RNase(10mg/ml, 1µl) mixture was added to the supernatant and incubated for 1hour at 37°C. Following this, 20% of PEG #8000 and 4ml of 2.5M NaCl was added. The tube was incubated on ice for 1 hour or overnight at 4°C. The mixture was centrifuged for 12 minutes at 6000rpm, and the supernatant was discarded.

The pellet was dissolved in 400µl of TM. 400µl of $CHCl_3$ was added and centrifuged at 5000rpm for 5minutes. The upper layer was transferred. Then 8µl of 10% SDS and 10µl of ProteaseK (10mg/ml) were added and incubated at 65°C for 15 minutes to 30 minutes. The mixture was extracted twice with phenol and once with chloroform. The DNA was then precipitated using 1/10 volume of 3M NaOAC and the equal volume of isopropanol and centrifuged for 30 minutes at 4°C. The DNA pellet was dissolved in 50µl of RNase/TE solution.

2.3.4 Cloning and Sequencing the bat *Hoxa2* region

The four positive clones were digested with *BamHI, EcoRI*, and *XbaI* in order to release the genomic fragments. Similar digestion pattern were visible among the four clones, presuming that they all contain the same genomic region, which has been confirmed by sequencing. These fragments were subcloned into Bluescript (Stratagene), and sequenced. Sequences were assembled by using Staden's Pregap4 and Gap4 software (Bonfield et al., 1995).

2.4 Transgenic Analysis

2.4.1 Reporter construct

The enhancer analysis was done in mouse and chicken. In mouse transgenic animals were generated in a modified version of BGZ40 (Yee and Rigby, 1993), containing a basal human beta-globin promoter, linked to the bacterial *beta-galactosidase* gene, an SV40 polyadenylation signal, and a multiple cloning site for easy removal of the vector sequences (Maconochie et al., 1999). For transgenic mice, construct were separated from vector sequences by electrophoresis and purified using MinElute (Qiagen), prior to microinjection. The production of transgenic mouse was carried out by pronuclear injection and *lacZ* reporter activity was detected as described.

For transgenic chicken, the entire plasmid was electroporated. DNA was diluted in water and Fast Green to a final concentration of 0.5-2μg/μl (Inoue and Krumlauf, 2001; Itasaki et al., 1999). The construct was electroporated in ovo, using the same detection system as in mouse. The embryos were electroporated in most cases between somite 6-8 and incubated for 18-24 hours after electroporation.

2.4.2 In ovo Electroporation

Fertilized chick eggs were obtained from Ozark hatcheries and were stored at 16°C for up to 7 days prior to horizontal incubation at 37°C in a humidified chamber. On the day of electroporation, a small hole was made at the sharp end of each egg. 2.0 - 2.5ml of albumin was removed through this hole using a 16 gauge syringe. The top of the shell was elliptically cut with scissors to open a window above the embryo. Black ink (Pelikan) diluted in Hanks' Balanced Salt Solution (GIBCO) was injected with a 26 gauge syringe under the embryos in order to increase contrast. A small tear in the vitelline membrane, directly above the injection site, was made using an electrolytically sharpened tungsten needle. A pulled glass capillary loaded with the DNA was placed through the hole in the vitelline membrane. The injection solution was then mouth pipetted into the neural tube.

Following injection of the DNA solution, two electrodes were placed on either side of the neural tube. The distance between the electrodes was approximately 4mm. Hanks' Balanced Salt Solution (GIBCO) was used to wet the electrodes. An electroporator, Cuy21 (Tokiwa Science, Japan), was used to generate and apply electric pulses. The conditions of the electric pulses were:

Voltage	18V
Pulse Length	50ms on
	950ms off
Pulse Number	5

Notes: V; volts, ms; milliseconds

Following electroporation, the window in the shell was sealed with transparent tape (Scotch), and the embryos were re-incubated at 37°C until the desired stages.

2.4.3 Production and analysis of transgenic mice

Mice were supplied by the Lab Animal Services Core Facility at the SIMR. The superovulation of the females was performed by the Lab Animal Services Core Facility at the SIMR. Briefly, female mice (3-4 weeks old) were superovulated by intraperitoneal (IP) injection of pregnant mare's serum (PMS), by IP injection of human chorionic gonadotrophin followed 46 hours later. Each female was mated with a stud-male and checked for a vaginal plug the following morning.

The oviducts were dissected out from plugged females, and the eggs were isolated in M2 medium. The cumulus cells were removed by incubation of the eggs in hyaluronidase for a few minutes. The eggs were washed twice in M2 medium, transferred into drops of KSOM medium under mineral oil, and equilibrated at 37°C in a humidified atmosphere with 5% CO_2.

30-40 eggs were transferred into a drop of M2 medium under oil on a Nikon inverted microscope and injected with the DNA construct with Leica-micromanipulators. The embryos were reimplanted into the oviducts of pseudopregnant plugged females. The pseudopregnant

female mice were anaesthetized by IP injection of 300-400µl (dependent on the weight) of 2.5% Avertin, and 10-15 eggs were transferred into each oviduct with a mouth pipette. The skin incision was closed with a wound clip, and mice were kept under a heat lamp to recover.

2.4.4 β-galactosidase assay

The embryos were dissected from the uterus in PBS and fixed for one hour in fixative at 4°C. They were washed once with PBS for two hours in wash solution at room temperature. The wash solution was then replaced with stain solution, and embryos were kept for up to 36 hours. After the staining was completed, the embryos were rinsed in PBS and stored long-term in PBS/0.01% sodium azide at 4°C.

2.5 Computer Comparisons

The global alignments were performed with MACAW (Schuler et al., 1991). Alignments were generated by using the *Hoxa2* and *Hoxb1* regions from the publicly available sequences and sequences from the chicken loci which I generated at SIMR. Local alignments were performed using Vector NTI's integrated ClustalW (Thompson et al., 1994) alignment program.

2.6 Electrophoretic mobility shift assay (EMSA)

A total of 35ml of 5% polyacrylamide gel was prepared containing the following reagents: 5% acrylamide (from 40% stock solution), 0.5x TBE, 2.5% Glycerol, 0.1% TEMED, 0.085% ammonium persulfate (APS) from a fresh 30% stock. For the antibody supershift assay, 3µg anti-Flag (M2) (Sigma) was added to the reaction mixture. Gels were imaged on a Typhoon 8700 (Amersham Biosciences).

Electrophoresis was performed in 1x TBE using vertical gel apparatus (SE 400, Amersham). The gel was equilibrated with running buffer for one hour at 150V at 4°C. The entire reaction volume was loaded onto the polyacrylamide gel and the gel was run for 3 hours at 150V at 4°C. Gels were imaged on a Typhoon 8700 (Amersham Biosciences).

2.7 Methods for Chapter 3

2.7.1 Isolation and Sequencing of the Chicken *Hoxa3/2* intergenic regions

The chicken 10.4kb genomic fragment, beginning in the intron between *Hoxa3* exons 1 and 2 and extending 3' of *Hoxa2*, was sequenced by subcloning and sequencing fragments and subsequently finished by primer walking (Tümpel et al., 2002). Sequences were assembled by using Staden's Pregap4 and Gap4 software (Bonfield et al., 1995).

2.7.2 Constructs for *in vivo* analysis in chick and mouse embryos

The chick intron (construct #3.5) was isolated by PCR with the primers GCT ACT AGT AGC CCA ACT TTC CCG AGT CG and AGC TCT AGA TTT ATA TCG GCC ATC GCG CG and cloned into Bluescript (Stratagene) using *SpeI* and *XbaI* and then cloned into BGZ40 using *SpeI* and *XbaI*. The bat intron sequence (construct #3.6) was cloned using PCR from the isolated phage fragments using the following primers: GCT TCT AGA GAC AAG CTT GGA ACT TTC CCT AAC TTG TG and TCG CCG CGG CCG GGA GCA AAA CTT TAT ATT A and cloned directly into BGZ40 using *SacII* and *NotI*. The mouse intron (construct #3.7) was cloned using PCR with the following primers: GCT TCT AGA GCT TCA ATA GTT TAA TAG TAG CG and GGA CTA GTC CTC CCC CGC CGC TGC CAT CAG C and subcloned into pGEM-T Easy (Promega) and cloned into BGZ40 using *SpeI*. The zebrafish intronic region (construct #3.8) was isolated from genomic DNA using PCR with the following primers: ACT TCG ACA GCG GCA ACC AC and GCC ACC ATC AGA AAT CTC GG and subcloned into pGEM-T Easy (Promega) and cloned into BGZ40 using *SacII* and *NotI*.

2.7.3 Mutagenesis

The following primers were used for the chicken r4 mutagenesis. PH1 mutagenesis (CGC CCC AAC GCG GGA TCG TTC GTT TGA GTT GGA GCT GAC C), PM mutagene-

sis (CCT AGA GCC GGG CTA TTT TAG AGT AAT GAA GAG TGA TAG ATT GC), PH2 mutagenesis (when the PH2 is also mutagenised) CCT AGA GCC GGG CTA TTT TAG AGT AAT GAA GAG TCG TAC G), PH2 (GCT ATT TGA CAG TAA TGA AGA GTC GTA CGT TGC TCC CGC TCA GCT CG), PH3 (GCT CAG CTC GGC GGC TCG TGC GTT AAT TAT CAA TCG CG). The following primers were used for the mouse mutagenesis: PM GCG GTC TAT CAC TCT TCA TTA CTC TAA AAA AGC CAA ACT CTA, PH2 GCT TTT TGA CAG TAA TGA AGA GTC GTA CGC CGC TCT TGC TCA GCT AAG CAGC, PH3 GCT CTT GCT CAG CTA AGC AGC TCG TGC GTT AAT TAT AAA TTG TGT TGT AGC.

2.7.4 Multimerisation Experiment

The annealed oligonucleotides were used for ligation with the appropriate vector. 2.PH GGC CGC AAG CTT AGT GAT AGA TTG CTA GTG ATA GAT TGC TAG TGA TAG TAG ATT GCT TCC CCG C; PM+2.PH GGG GAC AAT CTA TCA CTC TTC ATT ACT GTC ACA ATC TAT CAC TCTT CAT TAC TGT CAC AAT CTA TCA CTC TTC ATT ACT GTC AAA GCT TGC; 3.PH GGC CGC AAG CTT GGC TGA TGC ATT AAT TGG CTG ATG CAT TAA TTG GCT GAT GCA TTA ATT TCC CCG C; PH1 GGC CGC AAG CTT CGG GAT GAT GAT TTA TTT GAG CGG GAT TTA TTT GAG CGG GAT GAT TTA TTT GAG TCC CCG C; PM1+PH GGC CGC AAG CTT TGA CAG TAA TGA AGA GTG ATT TAT TTT GAC AGT AAT GAA GAG TGA TTT ATT TTG ACA GTA ATG AAG AGT GAT TTA TTT TCC CCG C; PM+PH3 GGC CGC AAG CTT TGA CAG TAA TGA AGA GTG ATG CAT TAT GAC AGT AAT GAA GAG TGA TGC ATT.

2.7.5 Electrophoretic mobility shift assay (EMSA)

Coupled TNT transcription/translation system (Promega) was used, labeled with ^{35}S-methionine (Amersham), to obtain Pbx1a, Hoxb1 and Prep1 protein. To ensure the appropriate size, proteins were visualized by SDS-PAGE followed by autoradiography. The EMSA conditions were the same as described in (Manzanares et al., 2001). Briefly, 2µl of reticulcyte lysate

containing the desired combinations of *in vitro* co-translated proteins were incubated in binding buffer (10mM Tris-Cl ph7.5, 75mM NaCL, 1mM EDTA, 6% glycerol, 3mM spermidine, 1mM DDT, 0.5mM PMSF, 1mg poly-dIdC, 40000 cpm ^{32}P-labeled oligonucleotide) in a total volume of 20µl. The reaction was incubated for 30 minutes on ice, the reaction were separated by 5% PAGE in 0.5x TBE. The oligonucleotide sequences used for the EMSA are the following: PH3 GCT CGG CGG CTG ATG CAT TAA TTA TCA ATC CG, PH1 CGC GGG ATG ATT TAT TTG AGT TGG AGC TGAC C, and Pbx/Meis PH2 GGC TAT TTG ACA GTA ATG AAG AGT GAT AGA TTG CTC CCGC. For the mutated versions of the Pbx/Meis and PH2, the following forward oligonucleotide sequences were used: Mut1 GGC TAT TTA AGA GTA ATG AAG AGT GAT AGA TTG CTC CCG C, Mut2 GGC TAT TTG ACA GTA ATG AAG AGT CGT ACG TTG CTC CCGC.

2.8 Methods for Chapter 4

2.8.1 Constructs for *in vivo* analysis in chick and mouse embryos

Construct #4.1 was obtained from the 10.4kb genomic fragment with a *BamHI*, *MluI*, digest, blunted and subcloned into *EcoRV* digested Bluescript vector (Stratagene). The insert digested with *HincII* and *XbaI* and then cloned into BGZ40 using *SpeI*, which was blunted followed by *XbaI* digest. For construct #4.2 construct #4.1 was used, digested with *EcoRI* and *BbsI*, blunted and then selfligated. This construct was then digested with *HincII* and *XbaI* and ligated into BGZ40 using *SpeI*, which was blunted followed by *XbaI* digest. Construct #4.3 was built by using PCR with the following primers: ACG CGT TTC AGC AGA ATG CG and GCA GGA CCC TGG GAG AGG AC and the product was subcloned into pGEM-T Easy (Promega) and then cloned into BGZ40 using *SacII* and *NotI*. Construct #4.4 was built by using PCR with the following primers: ACG CGT TTC AGC AGA ATG CG and AAA AAG TCA AAG CTG TCA GC and the product was subcloned into pGEM-T Easy (Promega) and then cloned into BGZ40 using *SacII* and *NotI*. Construct #4.5 was obtained by digesting construct #4.2 digested with *PstI* and then selfligated. This construct was then digested with *Hin-*

cII and *XbaI* and ligated into BGZ40 using *SpeI*, which was blunted followed by *XbaI* digest. Construct #4.6 was obtained by digesting construct #4.2 with *PstI,* the insert was then sub-cloned into Bluescript (Stratagene) which was digested with *PstI*. This construct was then digested with *HincII* and *XbaI* and ligated into BGZ40 using *SpeI*, which was blunted followed by *XbaI* digest.

The dog *Hoxa2* r2 regulatory region was cloned by PCR with dog (Labrador) genomic DNA using the following primers: GAA ATT TAA AAG CCT CGA AGA CTC and TGC TTT GTT TTG CTT TAA TGT TTT and the product was subcloned into pGEM-T Easy (Promega) and then cloned into BGZ40 using *SacII* and *NotI*. The mouse r2 enhancer was cloned using PCR with mouse genomic DNA using the following primers: GCT TCT AGA GAA ATT TAA AAA CCT GGA GGA C and AGC TCT AGA TTG TTT TTC AGG AAA ATC AC, the product was digested with *XbaI* and *SpeI* and subcloned into the Bluescript cloning vector (Stratagene). The insert was then inserted into BGZ40 after digestion with *XbaI* and *SpeI*. The zebrafish r2 enhancer of *Hoxa2* was cloned by PCR with zebrafish genomic DNA using the following primers: GAA AGA GAG GGT TAT CCA TT and GTT GGC TAT TTC TTT ATC CG and the product was subcloned into pGEM-T Easy (Promega) and then cloned into BGZ40 using *SacII* and *NotI*. The *Xenopus tropicalis* r2 enhancer was cloned by PCR with genomic DNA using the following primers: CTT GAA TAG TGT CTC TGG GG and CAA ATA CTC ATT TAT TCT ACC A, the product was subcloned into pGEM-T Easy (Promega) and then cloned into BGZ40 using *SacII* and *NotI*.

Construct #4.7 to #4.11 was cloned by PCR using for all of them the same reverse primer: TCG CCG CGG CTG AAG CTT CTG CAG GCA GGA ATC TGT GG and the following forward primers (for #4.7) GCT TCT AGA CAA TGG CGA ATC CCA AAG TT, (for #4.8) ATA AGA ATG CGG CCG CCC GTT TCG CCT TTA ACG AGC, (for #4.9)GCT TCT AGA AAA ATC TGA AAC ATT TTC AA, (for r#4.10) ATA AGA ATG CGG CCG CAC ACC AGT CAC CCA CTG TTC A, (for #4.11) GCT TCT AGA TCA ACA ATG GCC CAG AAC TG. Constructs #4.12 to #4.14 were cloned by PCR using for all of them the same

forward primer: GCT TCT AGA GAC AAG CTT ACG CGT TTC AGC AGA ATG CG and the following reverse primers (for #4.12) TCG CCG CGG CTC AGG ACT GTC ATT GTT GA, (for #4.13) TCG CCG CGG AGA GGC AGT TTT GAA CAG TG, (for #4.14) TCG CCG CGG TTC TCA TTG CTC GTT AAA GG. For all the listed constructs the PCR products were cut with *XbaI* and *SacII* and directly cloned into BGZ40.

For construct #4.36 the following primers were used: GGC CGC AAG CTT AAA CTG CCT CTC AAC AAT GGA AAC TGC CTC TCA ACA ATG GAA ACT GCC TCT CAA CAA TGG AAA CTG CCT CTC AAC AAT GGT CCC CGC and GGG GAC CAT TGT TGA GAG GCA GTT TCC ATT GTT GAG AGG CAG TTT CCA TTG AGA GGC AGT TTC CAT TGT TGA GAG GCA GTT TAA GCT TGC. For construct #4.37: GGC CGC AAG CTT CTC AAC AAT GAC AGT CCT GAC TCA ACA ATG ACA GTC CTG ACT CAA CAA TGA CAG TCC TGA CTC AAC AAT GAC AGT CCT GAT CCC CGC and GGG GAT CAG GAC TGT CAT TGT TGA GTC AGG ACT GTC ATT GTT GAG TCA GGA CTG TCA TTG TTG AGT CAG GAC TGT CAT TGT TGA GAA GCT TGC. For construct #4.38: CTG GAA AGC TTC TCA ACA ATG ACC TCA ACA ATG ACC TCA ACA ATG ACC TCA ACA ATG ACC CGC AND GGG TCA TTG TTG AGG TCA TTG TTG AGG TCA TTG TTG AGG TCA TTG TTG AGA AGC TTT. For construct #4.39: CTA GAA AGC TTA CAA TAC AAT ACA ATA CAA TCC GC and GGA TTG TAT TGT ATT GTA TTG TAA GCT TT. The annealed oligonucleotides were inserted directly into BGZ40 digested with *XbaI* and *SacII*.

2.8.2 Mutagenesis

The following primers were used for the chicken r2 mutagenesis (constructs #4.15 to #4.35). Only the forward oligonucleotide sequences are listed. All mutations were performed with construct #4.2 as a template. For the deletion constructs the following oligonucleotides sequences were used: (for Δ1) GCT CTC GCA GCA GCA GGC GCA GAA TGC GCC CCA AAG TTT CCC CGT TTC GCC TTT AAC G, (for Δ2) GGC GCA GAA TGC GCA CAA

TGG CGA ATC CCC GTT TCG CCT TTA ACG AGC AAT G, (for Δ2.1) GGC GCA GAA TGC GCA CAA TGG CGA ATA GTT TCC CCG TTT CGC CTT TAA CGA GCA ATG, (for Δ2.2) GGC GCA GAA TGC GCA CAA TGG CGA ATC CCA ACC CCG TTT CGC CTT TAA CGA GCA ATG, (for Δ3) GCA GAA TGC GCA CAA TGG CGA ATC CCA AAG TTT CCT TTA ACG AGC AAT GAG AAA AAT CTG AAA, (for Δ4) CGA ATC CCA AAG TTT CCC CGT TTC GGC AAT GAG AAA AAT CTG AAA CAT TTT CAA CAC C, (for Δ4.1) CGA ATC CCA AAG TTT CCC CGT TTC GCC AAC GAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (for Δ4.2) CGA ATC CCA AAG TTT CCC CGT TTC GCC TTT GCA ATG AGA AAA ATC TGA AAC ATT TTC A, (for Δ5) CCC AAA GTT TCC CCG TTT CGC CTT TAA CGA AAA TCT GAA ACA TTT TCA ACA CCA GTC ACC, (for Δ5.1) CCC AAA GTT TCC CCG TTT CGC CTT TAA CGA GAG AAA AAT CTG AAA CAT TTT CAA CAC CAG TCA CC, (for Δ5.2) CCC AAA GTT TCC CCG TTT CGC CTT TAA CGA GCA ATA AAT CTG AAA CAT TTT CAA CAC CAG TCA CC, (for Δ6) CCC CGT TTC GCC TTT AAC GAG CAA TGA GAC ACC AGT CAC CCA CTG TTC AAA ACT GCC TCT CAA CAA TGG C, (for Δ7) GCA ATG AGA AAA ATC TGA AAC ATT TTC AAA AAC TGC CTC TCA ACA ATG GCC CAG AAC TGC GC, (for the Δ8) CAT TTT CAA CAC CAG TCA CCC ACT GTT CAC CCA GAA CTG CGC AGC TGG CCT CAA CAA TGA CAG, (for Δ8.1) GAA ACA TTT TCA ACA CCA GTC ACC CAC TGT TCA TCA ACA ATG GCC CAG AAC TGC GCA GCT GGC, (for Δ8.2) CCA GTC ACC CAC TGT TCA AAA CTG CCT CCC CAG AAC TGC GCA GCT GGC CTC AAC AAT GAC AGT CC, (for Δ9) CCA CTG TTC AAA ACT GCC TCT CAA CAA TGG CTC AAC AAT GAC AGT CCT GAG GCC CTC GAG G, (for Δ10) CCT CTC AAC AAT GGC CCA GAA CTG CGC AGC TGG CGG CCC TCG AGG TCC CCT CTT TAC AGG, (for the Δ10.1) CCT CTC AAC AAT GGC CCA GAA CTG CGC AGC TGG CAC AGT CCT GAG GCC CTC GAG GTC CCC TCT TTA CAG G, (for Δ10.2) GCC CAG AAC TGC GCA GCT GGC CTC AAC AAT GGG CCC TCG AGG TCC CCT CTT TAC AGG ACT TTA ACG, (for Δ11) GCG CAG CTG GCC TCA ACA ATG ACA GTC CTG

ATA CAG GAC TTT AAC GTT TTC TCC ACA GAT TCC TGC.

For the site directed mutagenesis the following oligonucleotide sequences were used: (Δ1 TCC to TCG) GCA GCA GGC GCA GAA TGC GCA CAA TGG CGA ATC GCA AAG TTT CCC CGT TTC GCC TTT AAC GAG C,(Δ1 TCC to TCT) GCA GCA GGC GCA GAA TGC GCA CAA TGG CGA ATC TCA AAG TTT CCC CGT TTC GCC TTT AAC GAG C,(Δ1 TCC to TCA) GCA GCA GGC GCA GAA TGC GCA CAA TGG CGA ATC ACA AAG TTT CCC CGT TTC GCC TTT AAC GAG C, (Δ1 CAA to CAG) GCA GCA GGC GCA GAA TGC GCA CAA TGG CGA ATC CCA GAG TTT CCC CGT TTC GCC TTT AAC GAG C, (Δ2 CCT to CCG) CGA ATC CCA AAG TTT CCC CGT TTC GCC GTT AAC GAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ2 CCT to CCA) CGA ATC CCA AAG TTT CCC CGT TTC GCC ATT AAC GAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ2 CCT to CCC) CGA ATC CCA AAG TTT CCC CGT TTC GCC CTT AAC GAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ2 TTA to TTG) CGA ATC CCA AAG TTT CCC CGT TTC GCC TTT GAC GAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ2 ACG to ACA) CGA ATC CCA AAG TTT CCC CGT TTC GCC TTT AAC AAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ2 ACG to ACT) CGA ATC CCA AAG TTT CCC CGT TTC GCC TTT AAC TAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ2 ACG to ACC) CGA ATC CCA AAG TTT CCC CGT TTC GCC TTT AAC CAG CAA TGA GAA AAA TCT GAA ACA TTT TCA ACA CC, (Δ3 GAG to GAA) CCA AAG TTT CCC CGT TTC GCC TTT AAC GAG CAA TGA AAA AAA TCT GAA ACA TTT TCA ACA CCA GTC ACC, (Δ3 AAA to AAG) CCA AAG TTT CCC CGT TTC GCC TTT AAC GAG CAA TGA GAA GAA TCT GAA ACA TTT TCA ACA CCA GTC ACC, (Δ8 TCA to TCC) CCC ACT GTT CAA AAC TGC CTC TCC ACA ATG GCC CAG AAC TGC, (Δ8 TCA to TCT) CCC ACT GTT CAA AAC TGC CTC TCT ACA ATG GCC CAG AAC TGC, (Δ8 TCA to TCG) CCC ACT GTT CAA AAC TGC CTC TCG ACA ATG GCC CAG AAC TGC, (Δ8 ACA to ACC) CCC ACT GTT CAA AAC TGC CTC TCA ACC ATG GCC CAG AAC TGC,

(Δ8 ACA to ACG) CCC ACT GTT CAA AAC TGC CTC TCA ACG ATG GCC CAG AAC TGC, (Δ8 ACA to ACT) CCC ACT GTT CAA AAC TGC CTC TCA ACT ATG GCC CAG AAC TGC, (Δ8 ACAAT to AAAAT) CCC ACT GTT CAA AAC TGC CTC TCA AAA ATG GCC CAG AAC TGC, (Δ8 ACAAT to ACAAA) CCC ACT GTT CAA AAC TGC CTC TCA ACA AAG GCC CAG AAC TGC, (Δ8 ACAAT to ACAAG) CCC ACT GTT CAA AAC TGC CTC TCA ACA AGG GCC CAG AAC TGC, (Δ8 ACAAT to ACAGT), CCC ACT GTT CAA AAC TGC CTC TCA ACA GTG GCC CAG AAC TGC, (Δ8 ACAAT to AGAAT) CCC ACT GTT CAA AAC TGC CTC TCA AGA ATG GCC CAG AAC TGC, (Δ8 ACAAT to GCAAT) CCC ACT GTT CAA AAC TGC CTC TCA GCA ATG GCC CAG AAC TGC, (Δ8 CTCAA to CTTAA) CCC ACT GTT CAA AAC TGC CTC TTA ACA ATG GCC CAG AAC TGC, (Δ8 CTCAA to CGCAA) CCC ACT GTT CAA AAC TGC CTC GCA ACA ATG GCC CAG AAC TGC, (Δ8 TGGC to TCGC) CCC ACT GTT CAA AAC TGC CTC TCA ACA ATC GCC CAG AAC TGC, (Δ8 TGGC to TGTC) CCC ACT GTT CAA AAC TGC CTC TCA ACA ATG TC, (Δ10 AAC to AAT) GGC CCA GAA CTG CGC AGC TGG CCT CAA TAA TGA CAG TCC TGA GGC CCT CGA GG, (Δ10 AAT to AAC) GGC CCA GAA CTG CGC AGC TGG CCT CAA TAA CGA CAG TCC TGA GGC CCT CGA GG.

2.8.3 Computer Comparisons

For the codon usage analysis the following database was used: http://www.kazusa.or.jp/codon/.

2.9 Methods for Chapter 5

2.9.1 Constructs for *in vivo* analysis in chick and mouse embryos

2.9.1.1 *Hoxa2* r3/5 enhancer variants

The following primers were used for amplification of the fugu r3/5 modules: *Hoxa2(a)* TGG CTT AAT GCA AAC GCT AT and CCA TTA AGT TAA CAC TGA CAG ATA T, for

Hoxa2(b) TGC TGT AAT GCC AAA ACC TC and CCT GCC TCG CCT TCG TGC CG. The following primers were used for isolation of the medaka r3/5 modules: *Hoxa2(a)* GAC CTA TAT ATT TCA AAT GCA CAG G and GGA TTA AAC GTT GCA GTT CC, *Hoxa2(b)* CAG CAA AAG TCT CAG TTT ACA A and ATT GAT GGA GAA GAT CTG GG. The following oligonucleotides were used for introducing changes in the *Hoxa2(a)* modules (constructs #5.1-5). (construct #5.1) CCA CAG TTT GTA GGA GAG GCA GAG CTG CAC TGA AAG CCA ACA CCC ACT CAC CTC CTT GGA CAC AAA GCC TGT GCG TAA TTC, (construct #5.2) GCG ACA GCC TGG CTG TGA CTC CGA GCA GAA AAT GTG TCC TAT GCA CAC CTT GCT TGG TCC ACG TCC TGG CTG CAT TTG ATC CGG GGG AGA GTT AGA AGC, (construct #5.3) CTG AGT GGC CAG TGT TTC TCG CCG TTT CCA GGC ACT ATA TGA TCC CAG GGA GTG TTG GAT GCT TTA AAT GTG TTG CGA GGG CAC CGA GCT GTC AGA CC, (construct #5.4) CTG AGT GGC CAG TGT TTC TCG CCG TTT CCT GGC TGC ATT TGA TCC GGG GGA GAG TTA GAA GCC TTA AAT GTG TTC TTA GGG CAG GGA GCT GTC AGA CC, (construct #5.5) GCT GTC AGA CCT TTT GGC GAG TAA GAT TGA TCA CAC TCA GGG ACC GAG GTC TTT GTT TAG AGT CCG AGC AAC AAA CCT AGA GAG GCC TAC C, (Construct #5.6) CCA GCA CTC TTT GTT TGG TAT ATA AGC AAT AAA CAG CCA TCA GAT CCC ACC TTT CCT CTG CTT CCT CTC TCC CTC ACT TCC TTA CTC TCC C. For construct #5.7 the *Hoxa2 BglII* fragment was used as descript (Nonchev et al., 1996b). For construct #5.8 the sequence of construct #5.7 was used as a template for mutagenesis with the following oligonucleotide sequence: GGA ATA AAA GCA AGA AAA CTG GAA AAA CCC TTA CAT AAA ATA GCA TCT CTA TCT GCA AGG TAA TGC TCA GAG CTG G.

2.9.1.2 *Hoxa2* r4 enhancer variants

The fugu *(Fugu rubripes) Hoxa2* introns were isolated using PCR with genomic DNA with the following primers: *Hoxa2(a)* (construct #5.9) TTC CCC AAA AGG TGG GTG AT and ATC TCA GGC GAA CCT AGA AA and *Hoxa2(b)* (construct #5.10) TGA TGA TTA

ATA ACC TCT ATG TAA A and GAG CCT GAC TGG GAA GAT TT. The products were cloned into pGEM-T (Promega) and cloned into BGZ40 using *SacII* and *NotI*. The medaka *(Oryzias latipes) Hoxa2* introns were cloned from genomic DNA using PCR with the following primers: *Hoxa2(a)* GAT GCA GCC AGC TGC CCA CT and AGT TCT CAG CCG GCG GGA GC and *Hoxa2(b)* CGC TGC AGA CTG CGT CTA TA and TGG GTC TGC AAA.

For the fugu mutagenesis the following primers were used: PH1(a) to (b) change (construct #5.11) GCT TCG GGG CAA AAT GGG TAA TGA TTT AAG CAA GTT TGG TGG TGA TGA GAA ATG ATT TAT TCC; PH2(a) to (b) change (construct #5.12) CGC GTG TGA CAG TAA TGA AGA GTG ATA GAT TAC CGT TGC CGA GGA GGG CAG CTG G; PH3(a) to (b) change (construct #5.13) GGA GTG CCG TTG CCG AGG AGG GCA GCT GGC ATG TTC ATT AGT ATC CCA CAC TGG CCC TTG CTG C.

2.9.1.3 *Hoxa2* r2 enhancer variants

The fugu *Hoxa2(a)* r2 enhancers were isolated using PCR with the following primers: (construct #5.14) AGA TAC CTG CTC CTT TCA GA and CCA TAG GCC TAT TAT TGT AT; for *Hoxa2(b)* (construct #5.15) GGC AGT TCC AGC AAA ACC AAT and GCT TGC TTC TTT TGC AAA CA. The medaka *Hoxa2* r2 enhancers were isolated using PCR with the following primers: *Hoxa2(a)* GCT CCT TTC AGA ACA ACA CA and TAT TCA CAC ACA CTC AGC CA, for *Hoxa2(b)* GTT CCG ACC AAC ACA CAC CG and TTC AAT AGA ATT ATT TTG CT.

The changes in the r2 enhancer in constructs #5.18 to #5.22 were introduced using #5.17 as a template with the following oligonucleotides: (Construct #5.18) GCT CCA CCG CGG GGA AGA TAC CTG GCA GTT CCA GCA AAA CCA ATT CGG TTC GGC GCA GCA GCA CCA CAA TAG CG, (Construct #5.19) GCT CCT TTC AGA AAA AGT CTC ACA CTT CCC AGC AAT GTC TGA ACC ACA ATA GCG TGT CCA TGG GCT TTG CTG C, (Construct #5.20) CCT TTC AGA AAA AGT CTC ACG GTT CGG CGC AGC AGC ACG GGA ACA ATG GGG AGA GCC AAA GGC TTT GCT GCT GCG CCG CTG

AAC AGC, (Construct #5.21) GCA GCA GCA CCA CAA TAG CGT GTC CAT GGC CTG AGC GCA GCC CTT TTG AAC AGC AAT GAC AAA AAT CTG AAA CAT TTT CC, (Construct #5.22) GCA ATG ACA AAA ATC TGA AAC ATT TTC CAA ACG TGG CAC CCA CTG TTC CAA ACT GCG CGT CAA CAA TAG AGG CGG ACA ATC GTC ATT CCC C. Construct #5.20 was used as a template for construct #5.23. The following oligonucleotide sequence was used: GCA GCA GCA CGG GAA CAA TGG GGA GAG CCA AAG CCT GAG CGC AGC CCT TTT GAA CAG CAA TGA CAA AAA TCT GAA ACA TTT TCC. Construct #5.24 was made using construct #5.15 as a template with the following oligonucleotide sequence: GCA ATG TCT GAA CGG GAA CAA TGG GGA GAG CAT GGG CCT GAG CGC AGC CCT GCT GAA CAG CAA TGA GAA AAA TCT GAA ACA TTT TCC and construct #5.25 was made using construct #5.14 as template with the following oligonucleotide sequence: CCA CAA TAG CGT GTG CCA AAG GCT TTG CTG CTG CGC CTT TGA ACA GCA ATG ACA AAA ATC TGA AAC ATT TTC CCA ACC C.

For the fugu *Hoxa2(b)* and *(a)* construct (Construct #5.16) the second portion (fugu(a)) was cloned using PCR with the following primers: GCT GCA TGT CAA CAA TAG GC and CGG GGT ACC CCG ATA AGA ATG CGG CCG CTA AAC CAT AGG CCT ATT ATT GTA T, the product was digested with *HincII* and *KpnI* and cloned into Bluescript (Stratagene) digested with *HincII* and *KpnI*. The first portion (fugu(b)) was cloned using PCR with the following primers: TCT ATT GTT GAC GCG CAG TT and TCC CCG CGG GGA GGC AGT TCC AGC AAA ACC AA into pGEM-T Easy (Promega). The construct with the second portion (fugu(a)) was then digested with *SacII* and *HincII* and ligated with the insert of the first portion (fugu(b)), which was released from the vector by *SacII* and *HincII*. This construct was then digested with *SacII* and *NotI* and cloned into BGZ40 digested with *SacII* and *NotI*.

For the fugu *Hoxa2(a)* and *(b)* construct the second portion (fugu(b)) was cloned using PCR with the following primers: CGG GGT ACC CCG ATA AGA ATG CGG CCG CTA AAG CTT GCT TCT TTT GCA AAC A and CTG CGC GTC AAC AAT AGA GG, the

product was digested with *HincII* and *KpnI* and cloned into Bluescript (Stratagene) digested with *HincII* and *KpnI*. The first portion (fugu(a)) was cloned using PCR with the following primers: CTA TTG TTG ACA TGC AGC CT and TCC CCG CGG GGA AGA TAC CTG CTC CTT TCA GAA AA into pGEM-T Easy (Promega). The construct with the second portion (fugu(b)) was then digested with *SacII* and *HincII* and ligated with the insert of the first portion (fugu(a)), which was released from the vector by *SacII* and *HincII*. This construct was then digested with *SacII* and *NotI* and cloned into BGZ40 digested with *SacII* and *NotI*.

2.10 Methods for Chapter 6

2.10.1 Isolation and Sequencing of the Chicken *Hoxb1/2* intergenic region

The chicken *Hoxb1* genomic fragment was cloned from a cosmid containing the 3' region of the *Hoxb* cluster. *SmaII* fragments were subcloned, sequenced and subsequently the sequencing was finished by primer walking.

2.10.2 Construction of recombinant Baculovirus for FLAG mKrox20

The following primers were used to clone the mouse *Krox20*: TGC GAG GTC GAC ATG ATG ACC GCC AAG GCC GTA G and CGC GGT ACC GAG CTG GGC TCC ATC AAG GTC C. These primers generate a PCR product with a 5'*SalI* site and a 5'*KpnI* site, containing the mouse Krox20 gene including a short region of the 3'UTR. This product was cloned into a version of the pBacPAK8 vector which had been modified to include a region encoding a FLAG epitope tag between the *StuI* and *XhoI* sites in the multiple cloning sites, and in frame with the *XhoI* site. The *SalI/KpnI* digest of the PCR product was ligated into the *XhoI/KpnI* digested FLAG pBacPAK8 vector.

To generate recombinant baculoviruses Sf21 cells were co-transfected with the pBacPAK8 transfer vector containing FLAG-mKrox20 and linerized pBacPAK6 according to manufactures instructions. Briefly, 1.5ml suspension of $1x10^6$ exponentially growing Sf21

cells was added to a 35mm tissue culture dish and left for 30min to form a monolayer. The medium was aspirated and replaced twice with 2ml serum free medium. Meanwhile 86µl ddH2O, 5µl plasmid DNA (0.1mg/ml), 5µl pBacPAK6 linear viral DNA (Clontech) and 4µl Bacfectin transfection reagent (Clontech) were incubated at room temperature for 15 minutes. This was then added dropwise to the cells which were incubated at 27°C. Five hours later 1.5ml medium containing serum was added.

72 hours after co-transfection the supernatant containing recombinant viruses was harvested (~3ml) and centrifuged at 100xg for 5minutes in order to remove cell debris. 1ml of this supernatant was added to 100ml Sf21 cells in suspension at $5x10^5$ cells/ml in order to generate virus stocks. The supernatant from this flask was collected six days post infection by centrifugation at 100g for 5minutes and stored in the dark at 4°C.

2.10.3 Expression and Purification of recombinant Krox20 protein from Sf21 insect cells.

Sf21 cells were cultured at 27°C in Sf-900 II SFM (Invitrogen) with 10% fetal calf serum, 100units/ml penicillin, and 100µg/ml streptomycin. For the production of recombinant protein, 500ml of cells were cultured from a density of approximately $2x\ 10^5$cells/ml until they reached a density of $1x10^6$ cells/ml ($5x10^8$ cells). 20ml of recombinant baculovirus encoding the mouse Krox20 gene product with an N terminal FLAG tag were used to infect at a multiplicity of infection of approximately 10. Forty-eight hours after infection, cells were collected and lysed in ice cold buffer containing 50mM Hepes-NaOH (pH 7.9), 0.5M NaCl, 5mM $MgCl_2$, 0.2% Triton X-100, 20% (v/v) glycerol, 0.28µg/ml leupeptin, 1.4µg/ml pepstatinA, 0.17mg/ml phenylmethylsulfonyl fluoride, and 0.33 mg/ml benzamidine. After 30 minutes, lysates were centrifuged 100,000 x g for 30 minutes at 4°C and the supernatant collected for anti-FLAG immunoaffinity chromatography. Lysates were incubated in 1ml anti-Flag (M2)-agarose beads (Sigma) at least 12 hours at 4°C. The beads were washed three times with TBS and bound proteins were eluted by incubation at 4°C with 250µl TBS containing 10% glycerol

and 0.7mg/ml FLAG peptide. The eluate was collected and aliquots stored at -80°C.

2.10.4 Constructs for *in vivo* analysis in mouse embryos

The *Hoxb2* r3/5 enhancer was isolated from genomic mouse DNA with the following primers: TCT AGA GGA TCC CCA CTT TAA CAC CC and ACT AGT AAG CAG AGG GAA CAA CCT AA, and the product was cloned into the pGEM-T cloning vector (Promega). The insert was then isolated with *SpeI* and ligated into the *SpeI* site of BGZ40. The *Hoxb1* repressor region was cloned via PCR with the following primers: GGG CCC TTG AAG GGA GTG GA and AAA GAT CCT CTG CAC TTC TC. The product was cloned into pGEM-T cloning vector (Promega). The insert was then released with *SacII* and *NotI* and cloned into both BGZ40 and BGZ40 containing the *Hoxb2* enhancer.

To delete the *Hoxb1* Krox20 site, the following oligonucleotide sequence was used (only the forward sequence is listed): GGC CAG GCA GAC ACC CTG ACA AGT TAC AAA TGA TTG GAT TCT TGT CTT CAG AGT CTG GAG GAG G. To change the nucleotide from G to C at the third position, the following forward oligonucleotide sequence was used: CCT GAC AAG TTA CAA ATG AGA GTC GGT GTT GGA TTC TTG TCT TCA GAG TCT GG. To delete the RARE element, the following forward oligonucleotide sequence was used: CGC AGA GTG CCA CTG TTT ACG GAG ATC CCT CCC CCT GGA CTT GCC CTA GCT CAG GCC CCA GGC C.

2.10.5 Electrophoretic mobility shift assay (EMSA)

EMSA assay was performed using a modification of previously described methods with Cy5-labeled oligonucleotides (Sham et al., 1993). The probe was made by annealing two oligonucleotides, from which one was Cy5 labeled (Integrated DNA Technologies). The oligonucleotides were diluted to 100pm/µl and 50µl of each were mixed, incubated in a heating block at 95°C for 5 minutes and allowed to cool to room temperature. The annealed primers were then diluted 1:10 before being added to the reaction.

DNA binding reactions were set up in a 25µl volume and contained: 25mM HEPES pH

7.9, 50mM KCl, 10mM $MgCl_2$, 1mM EDTA, 10μ $ZnSO_4$, 1mM DTT, 8% glycerol, 5μg BSA, 100ng polydI-dC, and 5μl of Krox20 protein. The reaction was incubated for 10 minutes on ice. Following this, 1μl of probe was added, and the reaction was incubated for an additional 20 minutes on ice. The oligonucleotide sequences used for the probes were as follows:

Krox20-Hoxb1 AGG CCC CAG GCC TGT GGC CAG GCA GAC ACC CTG ACA AGT TAC AAA TGA GAG TGG GTG TTG GAT TCT TGT CTT CAG AGT CT

Krox20-Hoxb1(mut) AGG CCC CAG GCC TGT GGC CAG GCA GAC ACC CTG ACA AGT TAC AAA TGA GAG TCG GTG TTG GAT TCT TGT CTT CAG AGT CT

Krox20-Hoxa2 GAA GGC AAA AAG CTT TTT CAC CCA CGC AGC CTG ACA AAG CCC AAT GCT GTG GGC AGC CCT GCT TTC TAA CTT TCC TCT GT Krox20-Hoxa2(mut) GAA GGC AAA AAG CTT TTT CAG CCA CGC AGC CTG ACA AAG CCC AAT GCT GTC GGC AGC CCT GCT TTC TAA CTT TCC TCT GT

2.11 Methods for Chapter 7

2.11.1 Constructs for *in vivo* analysis in mouse embryos

Constructs #7.1 and #7.2 have been described previously (Marshall et al., 1994; Pöpperl et al., 1995). Construct #7.3 was isolated from a *Hoxb1 EcoRV* fragment (Marshall et al., 1994), digested with *HindIII*, the ends were filled in and then an additional *EcoRV* digest and cloned into BGZ40, which was digested with *NotI* and *XbaI* and the ends were filled in. Construct #7.4 was isolated from construct #7.3, digested with *ApaI*, blunted and digested with *HindIII* and cloned into Bluescript (Stratagene), which was digested with *SmaI* and *HindIII*. This construct was then digested with *NotI* and *HincII* and cloned into BGZ40 which was digested with *XbaI*, blunted and *NotI*. For construct #7.5, the *StuI-HindIII* fragment was isolated, blunted, and cloned into construct #7.1. This construct was then digested with *NotI* and *SacII* and the *ApaI* and *EcoRV* fragment of construct #7.3, cloned into Bluescript (Stratagene) was inserted into these two sites with *SacII* and *NotI*. For construct #7.6, construct #7.5 was digested with *AvrII* and *EcoRV* blunted and self-ligated. This construct was then digested with

NotI and *SacII* and cloned into construct #7.1. Construct #7.7 was derived from construct #7.5 by digestion with *AvrII* and *ApaI* and followed by self-ligation. This construct was then digested with *SacII* and *NotI* and cloned into construct #7.1.

Construct #7.8 was described previously (Maconochie et al., 1997). For construct #7.9, construct #7.8 was digested with *SacII* and *NotI*, and an insert of construct #7.5 was ligated into these sites. Construct #7.10 was described previously (Nonchev et al., 1996b). For construct #7.11, the insert of #7.5 was cloned into the *SacII* and *NotI* sites of construct #7.10. For construct #7.12, the insert of #7.6 was cloned into the *SacII* and *NotI* sites of construct #7.10. For construct #7.13, the insert of #7.7 was cloned into the *SacII* and *NotI* sites of construct #7.10.

For construct #7.14, a portion of the insert of #7.3 (CRI) was amplified using PCR with the primers ATG GTG ATA TCT TAC ATA AAA GG and GGG TTC AAG TCC TTG CTT GGA CA. The product was cloned into pGEM-T Easy (Promega), digested with *SacII* and *NotI*, and cloned into construct #7.10. For construct #7.15, the same PCR product was cloned into #7.2 using the restriction enzyme sites *SacII* and *NotI*. Construct #7.16 contains the CRII, which was amplified with the primers GAC AGT GGG TCC ATA CTG AAC C and CCT CCT AAC AGC TGG CCT AG, cloned into pGEM-T Easy (Promega), digested with *SacII* and *NotI*, and ligated into construct #7.10. Construct #7.17 contains CRIII, which was cloned using PCR with the following primers: CAA ACA AAC ACA GAC ATT GAC TGT and AGG TTG CAC TTT CTC CCA AA. This was cloned into pGEM-T Easy (Promega), and then digested with *SacII* and *NotI* and ligated into construct #7.10.

2.11.2 The targeted deletion of the *Hoxb1* repressor

The *Hoxb1* repressor targeting construct was constructed using homologous recombination. First the Neomycin resistance cassette was flanked with loxP sites using PCR with the following primers: ATA ACT TCG TAT AAT GTA TGC TAT ACG AAG TTA TAT TCA AAT ATG TAT CCG CTC ATG and ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG

TTA TTT TTA TTC TGT CTT TTT ATT GCC G. I cloned this construct into pGEM-T Easy (Promega). Next, this construct was flanked with sequences homologous to the flanking region of the *Hoxb1* repressor with the following primers: CTC TCC ATT TTC AAT GAA ATT TGC ATA TGA TAC AAT GCA TAT TAT GTA ATA TAT AAC TTC GTA TAA TGT ATG C and GGT ATC ACA GTG CTA TAC TGA GAT GAT TCC TAG ACT TTG CAT CTG GAA TAC GTA TAA CTT CGT ATA GCA TAC A. This construct was cloned into pGEM-T Easy (Promega).

A ~20kb genomic fragment that contained the *Hoxb1* coding sequence and additional upstream and downstream sequences was used (Marshall et al., 1994), as the basis for homologous recombination with the loxP and *Hoxb1* homology flanked neomycin cassette described above. The flanked neomycin cassette was released from the construct using *EcoRI* digestion, separated from its vector backbone by gel electrophoresis, and gel purification. The following protocol was used for the homologous recombination.

2.11.3 Homologous Recombination-mediated DNA engineering in *Escherichia coli*

2.11.3.1 Transformation of the *Hoxb1* construct into EL350 cells

A single colony of EL350 was incubated in 5ml LB and grown overnight at 32°C. The next day, 1ml of the overnight culture was used to inoculate 20ml LB and incubated at 32°C for 2-3 hours to reach the density of OD_{600}=0.5. The culture was then centrifuged at 0°C for 6 minutes at 5000rpm. The pellet was resuspended in 1ml of ice-cold water, transferred to a 1.5ml Eppendorf tube, and centrifuged using a bench-top centrifuge for 20 seconds at 13.000rpm. The pellet was washed three times with ice-cold water and then resuspended in 100µl of ice-cold water. The cells were transferred into an electroporation curvette, and 100ng of DNA was added to the cells. The electroporation was performed with a Biorad GenePulser electroporator at the following conditions: 1.8kV, 25µF, and 200ohms. Following this, 1ml of SOC medium was added to the cells and incubated in a 15ml tube for an hour. The cells were

spread onto LB plates containing ampicillin as a selection drug.

2.11.3.2 Using Homologous Recombination for inserting the neomycin resistance cassette into the *Hoxb1* construct

For inserting the flanked Neomycin cassette into the *Hoxb1* construct, the same protocol as described above used, with the following changes: after the cells transformed with the *Hoxb1* construct reached OD_{600}=0.5 to 0.8, 10ml of the culture was transferred to a new flask and incubated in a 42°C water bath for 15 minutes. Immediately after 42°C induction, the flask was placed into wet ice with constant shaking for 10-20 minutes. The cells were washed and resuspended as described above. For the electroporation, 100ng of DNA containing the Neomycin resistance cassette was added to the cells in the cuvette. The cells were then plated onto plates containing kanamycin as a selection drug. The colonies were then analyzed using restriction enzyme digest and sequencing.

The *Hoxb1* construct in which the Neomycin resistance casette had been inserted was then digested with *ApaI* and an 8.8kb fragment was released. This fragment was gel purified and cloned into *ApaI* digested Bluescript (Stratagene). This construct was then used for homologous recombination with G418-resistent ES cells.

2.11.4 Screening of G418-resistant cells

The initial screen for positive ES clones was PCR-based; two sets of nested primers were used for two separate PCR reactions. From the first PCR reaction, 1µl was added to the second reaction. The following sequences for the primers were used: (first set) GGG GGA GGC TAA CTG AAA CAC GGA and TGC GGA GGA AGC CAA AGC AGG T; (second set) GGA AGG AGA CAA TAC CGG AAG GAA CCC and CAG GTG GGA CGG GTC TGG GGT A. The entire reaction was then loaded onto a gel, and all the clones which gave a PCR product with a size of 2.7kb were then further analyzed using Southern gel transfer.

G418-resistent ES cells were allowed to grow to confluency in 24 well plates. Cells were rinsed with 1x PBS prior to adding 0.5ml/well digestion buffer (100mM Tris HCl (pH

8.0), 200mM NaCL, 5mM EDTA, 0.2% SDS, 0.2mg/ml proteinaseK). The digestion of the ES cells was allowed to proceed overnight at 37°C in a 5% CO2 incubator. Genomic DNA was recovered by adding 0.5ml of isopropanol per well, pipetting up and down, and transferring to 1.5ml Eppendorf tubes. These tubes were centrifuged for 15 minutes at 15000rpm at 4°C. The genomic DNA pellet was washed with 70% ethanol and dried for 15 minutes on the bench before being resuspended in 0.1ml 10mM Tris HCl (pH 8.0). Resuspension of the genomic DNA was facilitated by incubating the samples at 65°C.

2.11.5 *NcoI* Restriction Digestion of Genomic DNA and Southern Gel Transfer

20µl of each genomic DNA sample was digested overnight with 1µl of *NcoI* restriction endonuclease in the recommended buffer supplemented with 1µg/ml BSA, 1mM spermidine, and 1mM DTT. Samples were then loaded onto of 1% agarose gel and allowed to run for approximately 3 hours at 110 volts. The gel was then prepared for Southern transfer by submerging in the following solutions with gentle shaking: 10 minutes of depurination (0.2N HCl solution), 30 minutes of denaturation (1.5M NaCl, 0.5M NaOH), and 30 minutes of neutralization (3M NaCl, 0.5M TrisHCL, pH 7.5). Genomic DNA was transferred overnight in 20x SSC onto Hybond-N+ (Amersham Biosciences).

2.11.6 Hybridization, Post-Hybridization Washes and Exposure of Membranes to Film

The next day, the DNA was crosslinked to the membrane by incubation in an incubator at 80°C for 2 hours. I used two regions as probes, one 3' to the *Hoxb1* repressor region and the other one 5'. The 3' probe was cloned using PCR with the following primers: CCC ATT CTC TGA GCA AAC TAA A and CCT GGT TAG GTC TGT TAT CT; for the 5' probe, I used the following primers: GAG AGA GAG AAA CTT TTG ATC and TCA CTA AAG AGC ATA GCA GA.

Radiolabeled probe was prepared using the RediprimeII system (Amersham Biosciences) with 25ng of linerized DNA. Incorporated nucleotides were separated by using QIAquick Spin (Qiagen). The membrane was pre-hybridized in prewarmed Modified Church and Gilbert hybridization buffer (0.5M NaPO4 (pH 7.2), 7% SDS, 10mM EDTA) in an incubator at 65°C. Following this, radiolabeled probe was added to the buffer and incubated overnight at 65°C. The membrane was then rinsed once with post-hybridisation bufferI (2xSSC, 0.1% SDS) at room temperature, twice with post-hybridisation bufferI for 5 minutes at room temperature, twice with post-hybridisation bufferII (1xSSC, 0.1% SDS) for 10 minutes at room temperature, and twice with post-hybridisation bufferII for 10 minutes at 65°C. The membrane was placed in plastic wrap and taped onto an intensifier screen with a film on the top. The film was exposed for 3 days at -80°C before being processed.

2.12 Methods for Chapter 8

2.12.1 Methods used for the study: The distal enhancer contains a conserved Hox/Pbx site required for appropriate expression of the *RARβ* locus in the hindbrain.

2.12.1.1 Recombinant proteins and EMSA

Purified GST-Hoxb4 fusion was produced as previously described (Gould et al., 1997). Electrophoretic mobility shift assays were performed using a modification of previously described methods (Gould et al., 1997), with Cy5 labeled oligonucleotides. The oligonucleotides were as follows: HS-1+HS2(Hoxb4) GAG AAT TAT ACA GAA AAC CAT TAA TCA CTT; PH Site (RARβ) TTT GAG GAG CAG GGT GAT AAA TAA TGG GGC TTT TCC A; Mutant PH site (RARB) TTT GAGG AGC AGG GTG GGC CCG CCG GGG GCT TTT CCA; Mutated oligonucleotides for the HS-1+HS2 have been described (Gould et al., 1997). Gels were imaged on a Typhoon 8700 (Amersham Biosciences).

2.12.1.2 Plasmid constructions for *in vivo* analysis in chick embryos

The 2.3 kb *NheI* fragment was isolated using PCR with mouse genomic DNA. The sequences for the PCR primers are: AGA ATG TGT GTG CTG ACT CTG C and AAG CAG TCT TAC CAG GAG GG. The product was subcloned into pGEM-T Easy Vector Systems (Promega); it was then cloned into the *SacII* and *NotI* sites of the BGZ40 vector. The PH deletion was performed using the following oligonucleotide sequence: CCT TGT GAA GTC CCC TTT GAG GAG CAG GGT GGG GCT TTT CCA ATT GTT ATT TGC CAA AAG G.

To create the multimerized construct (3xPH site), the following complementary oligonucleotides were subcloned: GGC CGC AAG CTT GAG CAG GGT GAT AAA TAA TGG GGC TTG AGC AGG GTG ATA AAT AAT GGG GCT TGA GCA GGG TGA TAA ATA ATG GGG CTT CCG C and GGA AGC CCC ATT ATT TAT CAC CCT GCT CAA GCC CCA TTA TTT ATC ACC CTG CTC AAG CCC CAT TAT TTA TCA CCC TGC TCA AGC TTG C into *NotI/SacII* digested pBSKS and then this sequence was moved into BGZ40. All constructs were sequenced to confirm their accuracy.

2.12.1.3 Sequence Analysis

RARβ intron sizes were determined by BLAST-like Alignment Tool (BLAT, Kent, 2002) comparisons of known RARβ mRNAs (Mendelsohn et al., 1994a; Zelent et al., 1991) with the UCSC Genome Browser Databases (Karolchik et al., 2003, see http://genome.ucsc.edu) using the following genome builds: human May 2004 (hg17), chimp Nov 2003 (pan Tro1), mouse May 2004 (mm5), rat Jun 2003 (rn3) and dog Jul 2004 (canFam1). The distance between the HP site of the distal enhancer and the ATG of the first coding exon (E2) of the RARβ1 mRNA isoform is at least 241kb in chimp and 215kb in rat, although gaps in these genome builds indicate that these distances may be larger. There are no gaps in human (240kb interval) and only a small number of gaps in dog (232kb interval) and mouse (~284kb). The position of the P1 promoter is not yet clear in rat, chimp and dog as 5' untranslated exon(s) of RARβ1 have yet to be identified. For human, however, Toulouse et al

(1996) have reported the 5' end of the RARβ1 isoform (Genbank Acc # U49855) which our BLAT comparisons show to be encoded by 6 non-coding exons, 3 of which lie upstream of the HP site. Multispecies HP site comparisons used MULTIZ (Blanchette et al., 2004).

2.12.2 Methods used for the study: The *Hoxb1* enhancer and control of rhombomere 4 expression: Complex interplay between PREP1-PBX1-HOXB1 binding sites.

2.12.2.1 Constructs for *in vivo* analysis in chick and mouse embryos

Subcloned fragments of the regulatory regions of the mouse *Hoxb* gene were cloned into the BGZ40 vector (Maconochie et al., 1997). A 622 bp genomic fragment which contains the 331 bp *StuI-HindIII* fragment of *Hoxb1* spanning the R1, R2 and R3 Pbx-Hox (PH) sites and the PM1 and PM2 Prep/Meis (PM) sites and functions as an r4 autoregulatory enhancer (Pöpperl et al., 1995), was used as a control (WT) for comparison with variants in which PM1 and or PM2 were mutated. The 622bp enhancer was isolated by PCR from mouse genomic DNA using CG CGG CTA GTC ATC CTT TTG TCC CAA GA and CCG CGG TCT TGC CCT ACA ACC TTT CG. The fragment was cloned into pGEM-T (Promega) and then transferred as a *SacII* fragment into the BGZ40 vector (Maconochie et al., 1997). The PM1, PM2 and PM1+PM2 sites were mutated (nucleotides substitution in PM1, underlined, and deletion in PM2, compare with the sequence in Figure 8.3B) using the following oligonucleotides: PM1, GGG CTC AGA GTG ATT GAA GTG TCT TGC TGT AGC TAA TGA TTG GGG GGT GAT GGA TGG; PM2, GGG GGG TGA TGG ATG GGC GCT GG GG AAA CTC TGG CCC GCT TAG CCC ATT GGC C; PM1+PM2, GGC TCA GAG TGA TTG AAG TGT CTT GCT GTA GCT AAT GAT TGG GGG GTG ATG GAT GGG CGC TGG G GAA ACT CTG GCC CGC TTA GCC CAT TGG CCT GGG. All constructs were sequenced to verify the mutations. The 622bp WT and variant PM1 and/or PM2 constructs were assayed in chick and mouse embryos. In mouse experiments, purified insert sequences were isolated from the

vector backbone by digestion with *ScaI* and *XhoI* and gel electrophoresis; the entire plasmid was used for chick electroporation.

Chapter 3

Regulation of *Hoxa2* in rhombomere 4

3.1 *Hoxa2* regulation and function during hindbrain development

In mouse, *Hoxa2* is expressed in early stages up to an r2/r3 boundary at low levels and then becomes progressively upregulated, specifically in r3 and r5, through *trans*-activation by Krox20 (Hunt et al., 1991; Krumlauf, 1993; Nonchev et al., 1996b; Wilkinson et al., 1989) (Figure 3.1). In contrast, *Hoxa2* is strongly expressed from r2-r6 in the chick hindbrain from the early stages, and there is no apparent upregulation specifically in r3 and r5, nor lower levels in r4, as seen in the mouse (see Figure 3.1 and Prince and Lumsden, 1994). The expression in r2 is also very strong and occurs throughout all cells, not just dorsally, as detected in the r2 of mouse embryos.

Hoxa2 is expressed in NCC derived from r4 migrating into the second BA but is not expressed in NCC derived from r2 (Prince and Lumsden, 1994). *Hoxa2* has been shown in various model systems to play a crucial role during craniofacial development by its ability to determine the morphological fate of second BA NCC (Gendron-Maguire et al., 1993; Grammatopoulos et al., 2000; Pasqualetti et al., 2000; Rijli et al., 1993). In *Hoxa2* null mutants, the second BA derived structures are transformed to a first arch fate (Gendron-Maguire et al., 1993; Rijli et al., 1993). In contrast, if *Hoxa2* is overexpressed in the first BA, its derivatives develop second arch-like structures (Grammatopoulos et al., 2000; Pasqualetti et al., 2000). The dramatic effect of the *Hoxa2* gain- and loss-of- function experiments has been explained by its ability to limit bone formation; in the second BA, chondrogenesis occurs only in regions which do not express any *Hoxa2* (Kanzler et al., 1998). It has been shown that Hoxa2 represses the expression of *Sox9* and *Cbfa1* in the second BA, which are the earliest markers for chondrogenic and osteogenic pathways (Couly et al., 2002).

Several *cis*-regulatory elements of *Hoxa2* have been identified; r3/5 expression is mediated by a complex regulatory region located in the intergenic region of *Hoxa2/a3*, which in-

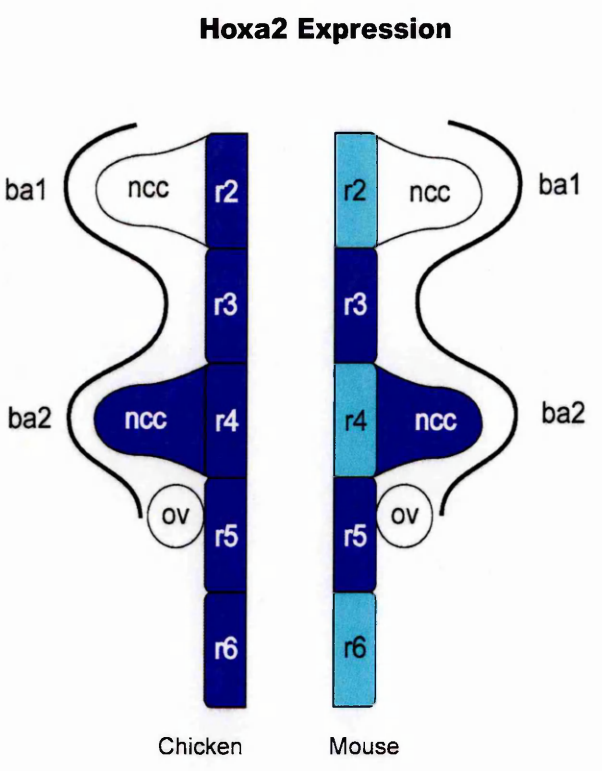


Figure 3.1. Expression domains of mouse and chick *Hoxa2* in the hindbrain.
Domains of expression of *Hoxa2* as determined by *in situ* hybridization are compared between chick and mouse. Intensity of blue (*Hoxa2*) correlates with the level of expression in the rhombomeres (r2-r6) and neural crest cells (NCC). The position of the otic vesicle (ov) and branchial arches (BA1, BA2) are also noted. NCC, neural crest cells.

cludes binding sites for the zinc-finger transcription factor, Krox20 (Nonchev et al., 1996b). Interdigitated in the same control region are several elements controlling *Hoxa2* expression in NCC. One of these elements has been shown to contain the binding site for the transcription factor AP2. However, it is unknown how *Hoxa2* expression is regulated in r2 and r4.

I will discuss in my thesis the basis of the regulation in r4 (Chapter 3) and in r2 (Chapter 4). In the next Chapter (Chapter 5), I will present data explaining the basis of the differential expression of the two fugu co-paralogous genes *Hoxa2(a)* and *Hoxa2(b)*. The work in Chapter 3 was performed in collaboration with Dr. E. Ferretti.

3.2 Expression of *Hoxa2* in Rhombomere 4 is regulated by a Conserved Cross-regulatory Mechanism

Hox genes of the paralog groups I-IV are expressed in the hindbrain in rhombomere restricted domains (Lumsden and Krumlauf, 1996). In r4, Hoxb1 plays an important role in patterning and in establishing segmental identity (Gavalas et al., 1998; Studer et al., 1996; Studer et al., 1998). The expression of *Hoxb1* is initiated by retinoids and maintained by a conserved auto- and cross-regulatory mechanism through interactions with Hoxa1 and Hoxb2 (Gavalas et al., 2003; Marshall et al., 1994; Pöpperl et al., 1995; Studer et al., 1998). The binding and transcriptional specificity of Hox factors is achieved by heterodimerization with the homeodomain protein Pbx, which is the murine homolog of the *Drosophila* extradenticle (exd) (Ferretti et al.,

1999; Ferretti et al., 2000; Mann and Chan, 1996). Other important cofactors for mediating specificity of Hox binding and transcriptional activity are Prep and Meis proteins, which are the murine homologs of *Drosophila* homothorax (hth) (Chang et al., 1997). These factors interact with Hox proteins, form ternary complexes, and regulate transcription of their target genes.

Hoxb2, the paralogous member of *Hoxa2*, is also expressed in r4, and it has been shown that this expression is mediated by Hoxb1 and its co-factors Prep/Pbx. The expression is regulated by a highly conserved module located in the intergenic region of *Hoxb2/b3* adjacent to the r3/5 enhancer elements.

In this Chapter, I show that the location of the *Hoxa2* r4 enhancer is different than that of *Hoxb2*, but the same *trans*-acting factors are involved. I used comparative sequence analysis and identified a highly conserved region in the *Hoxa2* intron. I show that this region contains Hoxb1/Pbx and Prep/Meis binding sites. Deletion analysis show that these sites are crucial for r4 enhancer activity, and gain-of-function experiments show that Hoxb1 is sufficient to *trans*-activate the *Hoxa2* r4 enhancer in the hindbrain.

3.3 Results

3.3.1 Alignments of the *Hoxa2/a3* intergenic region and *Hoxa2* coding region and the identification of the *Hoxa2* r4 enhancer

The intergenic genomic sequences of *Hoxa2/a3* and coding region of *Hoxa2* in chicken and bat were cloned, sequenced and compared with other publicly available vertebrate genomes including human, mouse, rat, zebrafish, fugu, and shark. One previously described, highly conserved region was present in the intergenic region, which encompasses the *Hoxa2* r3/5 and NCC enhancers (Maconochie et al., 2001; Nonchev et al., 1996b; Tümpel et al., 2002; Zhang et al., 2000). However, I was unable to identify any other conserved elements in the intergenic region that might direct *Hoxa2* expression in other rhombomeres in particularly in r4 (Figure 3.2A).

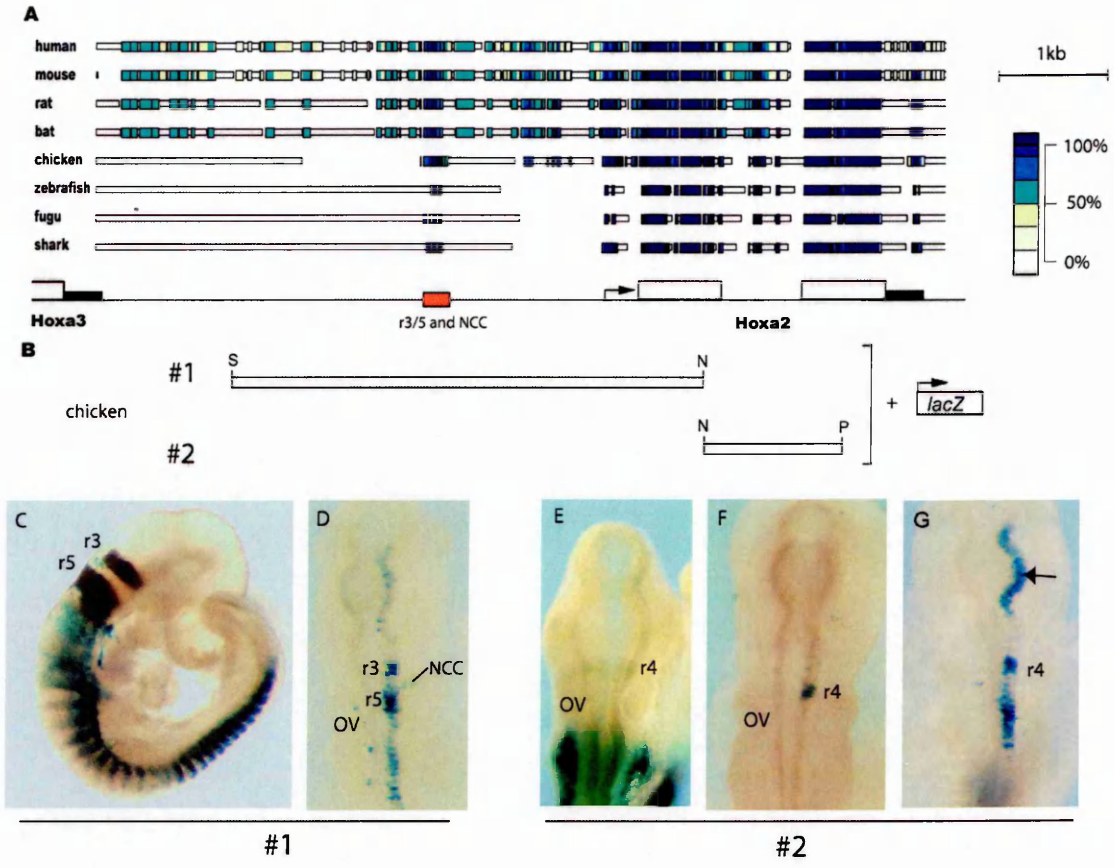


Figure 3.2. Alignment of the *Hoxa2* regulatory region and regulatory analysis of the chick *Hoxa2* intergenic and coding region.
A shows a MACAW alignment of the *Hoxa2*/a3 intergenic and *Hoxa2* coding region among different species. The color code on the left shows the correlations between the color in the alignment and the degree of conservation (0-100%) among the analyzed genomic fragments. The red box represents the location of the previously identified r3/5 enhancer (Nonchev et al., 1996b). **B** delineates constructs #1 and #2 used in the transgenic mouse and chicken elctroporation experiments. Construct #1 is a chicken *SpeI-NotI* (S and N) fragment spanning the *Hoxa3/2* intergenic region; the *NotI-PstI* (N and P) fragment (#2) includes the *Hoxa2* intron region. **(C)** Lateral view of transgenic mouse carrying construct #1 and **(E)** dorsal view of a mouse embryo carrying construct #2. Note the strong expression in the somites and weak expression in r4. **D** shows dorsal view of chicken embryos electroporated with constructs #1, and **F** and **G** show chick embryos electroporated with construct #2. Note the variation of expression in the neural tube between these two embryos, including ectopic expression in the forebrain (arrow) S, *SpeI*; N, *NotI*; P, *PstI;* r3, rhombomere 3; r5, rhombomere 5; r4, rhombomere 4; ov, otic vesicle; ncc, neural crest cells

The r4 module of the *Hoxa2* paralogous member, *Hoxb2*, is embedded adjacent to the r3/5 module in the intergenic region. To determine whether this is also the case for the *Hoxa2* regulatory network or whether there were any other elements present in the chicken intergenic *Hoxa2*/a*3* region unrecognized by the sequence comparison approach, I cloned this region (Figure 3.2B, construct #1) upstream of the *lacZ* reporter gene and tested it in chick electroporation and transgenic mouse assays (Figures 3.2 C, D). In both of these assays, the fragment efficiently directed expression in rhombomeres 3 and 5 (r3/5). In addition, the electroporated chick embryos also showed reporter expression in NCC (Figure 3.2D), and somatic expression

of the reporter gene was observed in the transgenic mouse embryos (Figure 3.2C). Therefore, no other measurable rhombomeric enhancer elements were present in this region, confirming the sequence comparison analyses (Figure 3.2A).

I extended the comparisons and included the region of the *Hoxa2* coding sequence. As shown in Figure 3.2A, a second conserved region lays in the intron of *Hoxa2*. A construct encompassing the intron was tested upstream of the *lacZ* reporter gene in transgenic mice. The reporter was weakly expressed in r4 and strongly expressed in the lateral mesoderm and somites (Figure 3.2E); the same construct, when electroporated into chick embryos, directed strong r4 expression (Figures 3.2F and G). In these examples, both chick embryos showed strong r4 expression, but the embryo in Figure 3.2G showed some additional neural expression and represents the variation in this assay.

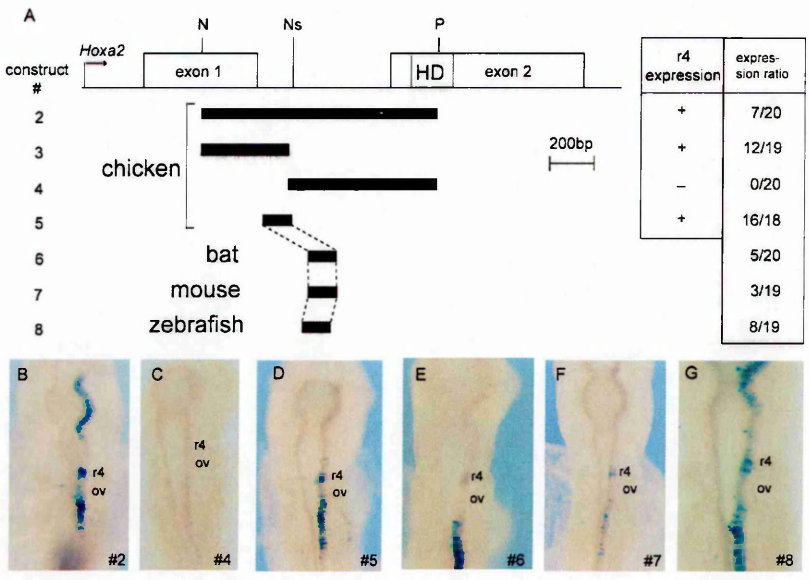


Figure 3.3. Identification of the r4 enhancer.
A illustrates the constructs used to identify the r4 enhancer found to be located in the *Hoxa2* intron. Constructs #6-8 are derived from the orthologous region of bat, mouse, and zebrafish *Hoxa2* introns. (**B-G**) Dorsal views of representative chick embryos electroporated with constructs #3-8. Ns, *NsiI*; N, *NotI*; P, *PstI;* r4, rhombomere 4; ov, otic vesicle.

In order to narrow down the r4 enhancer in the intron region, I tested various constructs in the chick electroporation assay system (Figures 3.3). I subdivided the chicken *Hoxa2* intron into two regions (Construct #3 and #4) and evaluated them for r4 enhancer activity. The construct (#3) containing the 5' region of the intron showed enhancer activity in rhombomere 4, whereas the construct (#4) with the 3' region of the intron did not exhibit any reporter activity. This is consistent with the sequence comparison, which showed that the highest conservation is in the 5' end of the intron (see Figure 3.2A). I next asked whether the highly conserved region alone is sufficient to drive r4 reporter expression. I isolated the conserved region, linked

it to a *lacZ* reporter construct (construct #5), and assayed the construct in chick electroporation; in the vast majority of the electroporated embryos, I observed r4 restricted reporter staining (Figure 3.3D).

I also cloned the homologous region from different species, including bat, mouse, and zebrafish; in each case, the element was able to direct r4 specific reporter staining in electroporated chick embryos, although the expression was weaker as compared to the chicken r4 en-

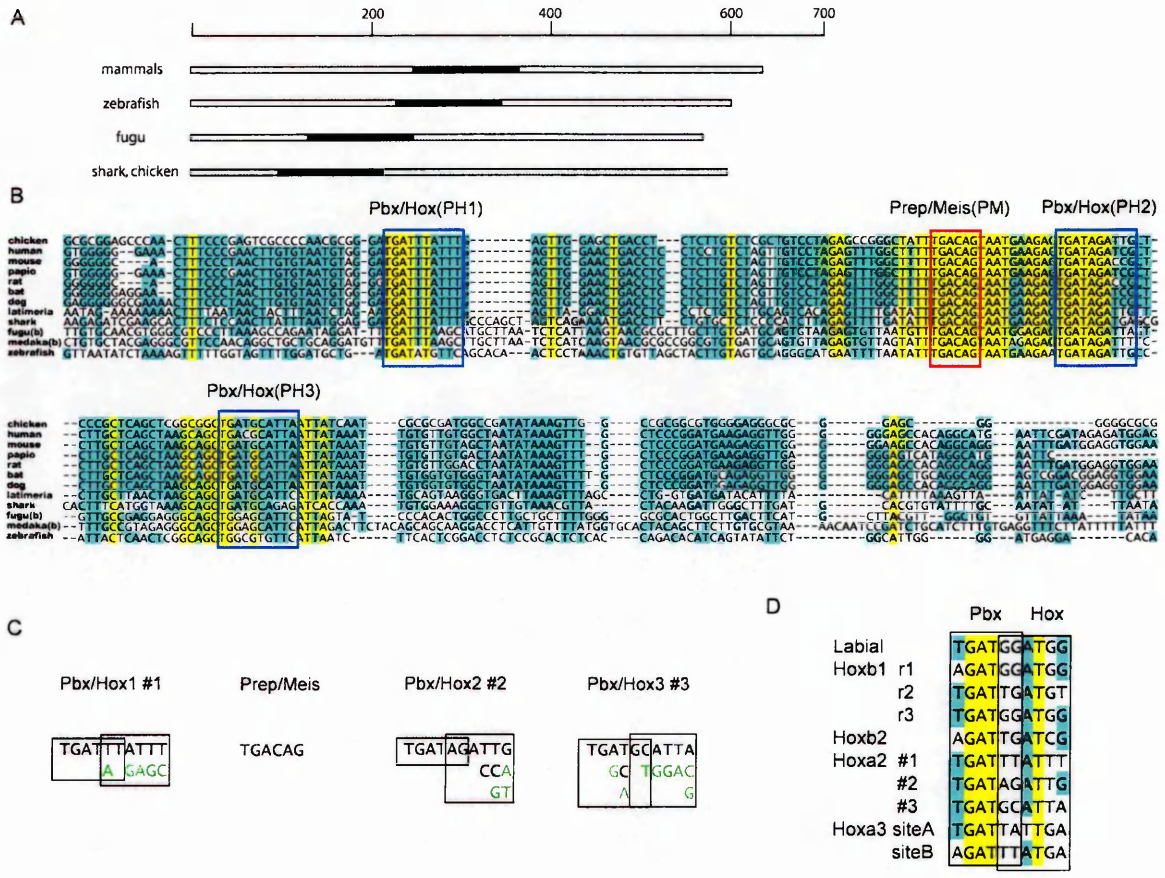


Figure 3.4. Organization, conservation, and alignment of *Hoxa2* r4 control region directing r4 expression. **A** shows the relative location of the highly conserved region (black boxes) within the intron in different species. Local alignment is shown in **B**, with potential PH sites (1-3) boxed in blue and the Prep/Meis element boxed in red. The colors within the sequence alignment indicate the degree of identity: yellow indicates complete conservation of sequence; blue indicates majority of species in which the bases are conserved. Consensus sequence of each site is shown in figure **C**. **(D)** Various Hox/Pbx sites found in *Hox* target genes are listed and compared with those identified in *Hoxa2*.

hancer (see Figures 3.3E, F, and G). This suggests that there are slight differences in the *Hoxa2* r4 elements directing species-specific expression (see also Figure 3.1).

3.3.2 Identification of Hox, Pbx, and Prep binding sites in the r4 enhancer of *Hoxa2*

The position of each highly conserved region within the intron differs to some degree among species (Figure 3.4A). In mammals, the highly conserved region is located in the middle of the intron, whereas in zebrafish, fugu, shark, and chicken, the highly conserved regions are located closer to the beginning of the intron. Thus during evolution of *Hoxa2*, the r4 element has been conserved, but insertions/deletions surrounding it are likely to have occurred and shifted the relative position within the intron.

The regulation of other genes expressed in r4, including *Hoxb1, EphA2,* and *Hoxb2*, is dependend upon elements which bind Hox/Pbx and Prep/Meis (Chen and Ruley, 1998; Maconochie et al., 1997; Pöpperl et al., 1995). I therefore examined the *Hoxa2* highly conserved region directing r4 expression for potential transcription factor binding sites. Careful inspection of the aligned sequences of 12 different species highlighted blocks of high conservation which included similar motifs, consisting of three potential Hox/Pbx (PH1-3) sites (blue boxes) and one potential Prep/Meis (PM) site (red box) (Figure 3.4B).

The putative Prep/Meis site shows a very high degree of conservation, with no variation among the species (TGACAG) (Figure 3.4C). This is the exact sequence of the motif characterized in the r4 enhancer of the *Hoxa2* paralog member, *Hoxb2* (Maconochie et al., 1997), although the sequence is complimentary. The characterized TGACAA PM site of the *Hoxb1* and *Hox*a3, is very similar, varying only in the last nucleotide (Manzanares et al., 2001; Pöpperl et al., 1995). However, the putatitive *Hoxa2* PH sites show more variation (Figure 3.4C); these motifs are highly conserved among amniotes, but the variation increased when the sequences of fishes were added to the alignment (see bases in green font which reflects increased variation due to the inclusion of fish genomic sequences in Figure 3.4C). I did not include the sequences of medaka and fugu *Hoxa2(a)* in the final alignment, since it has been shown that they are not active in r4 (see also chapter 5 and Amores et al., 2004).

PH binding sites have been identified in the *Hoxb1* r4 enhancer (autoregulatory ele-

ment), in the *Hoxb2* r4 enhancer, and in the *Hoxa3* r5/6 enhancer. When comparing the mouse sequences of these sites with the three putative mouse PH sites of the *Hoxa2* r4 enhancer, it is possible to define a similar pattern in a consensus sequence (Figure 3.4D). The Hox and Pbx binding sites overlap, leaving a core in the middle with high variation among different mouse *Hox* gene enhancers (Figures 3.4 C and D). The three putative PH sites show variation in their core sequences. The second putative *Hoxa2* bipartite site, PH2, has GG or AG in its core, while the third bipartite site contains a GC or GT core. These two cores resemble the *labial/Hoxb1/Hoxb2* cores, which contain the sequences GG, TG, or AG (Maconochie et al., 1997; Manzanares et al., 2001; Pöpperl et al., 1995). The first putative bipartite Hox/Pbx (PH1) site of *Hoxa2* is different from these sites, since it contains a central dinucleotide, TT, resembling the core sequences identified in the bipartite PH site responsible for the regulation of the *Hox* paralogous groups 3-10 (Manzanares et al., 2001).

The comparison among various mouse PH sites of different enhancer regions shows that there are two blocks of conserved sequence. The first one is present in the Pbx binding site and encompasses a highly conserved GAT sequence. The other one is located in the Hox binding site with the sequence AT, although the *Hoxa3* siteA varies at this positions with the TT sequence (Figure 3.4D). These two conserved regions are also conserved within the three mouse *Hoxa2* PH sites (Figures 3.4D).

This comparison between the sequences of the three *Hoxa2* PH sites and of other known PH sites suggests that the identified r4 elements of *Hoxa2* display a sequence pattern very similar to already described PH binding site of other *Hox* control elements.

3.3.3 *In vitro* binding analysis of Hoxb1, Prep, and Pbx to the r4 enhancer

The sequence comparison of the PH sites of the *Hoxa2* r4 enhancer suggests that they are PH binding sites. Since it has been shown that ectopic *Hoxb1a* expression in zebrafish can *trans*-activate *Hoxa2* expression (Hunter and Prince, 2002) and, further, that Hoxb1 plays a

major role in r4 (Studer et al., 1996), I tested whether Hoxb1 and its co-factors are able to bind to the *Hoxa2* r4 enhancer element *in vitro*. I used an established electrophoretic mobility shift assay (EMSA) to evaluate the binding activity of these elements (Ferretti et al., 2000; Manzanares et al., 2001). I used labeled double stranded oligonucleotides spanning the proposed elements to test their ability to bind Prep, Pbx, and Hoxb1 (Figure 3.5). When the oligonucleotides were incubated in the presence of Pbx1a and Hoxb1, a dimer complex was formed (lane 4), which is also the case when incubated with Prep and Pbx1 (lane 5). When all three proteins are present, a ternary complex is visible (lane 6). This is an important finding, since it has been shown that ternary complex formation with these three proteins is capable of increasing transcription (Berthelsen et al., 1998). These results were similar to those previously reported for the *Hoxb2* r4 enhancer (control in Figure 3A, first three lanes, and Ferretti et al., 2000; Maconochie et al., 1997; Manzanares et al., 2001).

I then asked whether mutation of the *Hoxa2* PH2 and PM sites has an effect on complex formation (lanes 8 -15). Oligonucleotides in which the Prep/Meis site were mutated were still able to form expected dimer complexes between Hoxb1 and Pbx1a (lane 8) and Pbx1a and Prep (lane 9), but failed to form a ternary complex when all three proteins are present (lane 10). Oligonucleotides in which the sequence of the PH2 site was changed were unable to mediate complex formation even in the presence of all three proteins. Therefore, the presence of both sites is necessary for ternary complex formation.

The first and third PH sites did not form dimer complexes with Pbx and Hoxb1 (data not shown). I therefore decided to add the PM sites adjacent to the first and the third PH sites, since I showed that the presence of both sites, PM+PH2, are important for ternary complex formation (see Figure 3.5A). When I added either of these oligonucleotides, PM+PH1 and PM+PH3, I observed strong dimer formation between both PM and Hoxb1/Pbx1a (lanes 21 and 27). Also both oligonucleotides formed ternary complexes between Prep/Pbx and Hoxb1 (lanes 23 and 29). This shows that dimer formation of the PH2 is independent of the PM site, whereas ternary complex formation is formed only in additional presence of the PM site.

The first and third PH sites were unable to mediate dimer formation, whereas in the presence of the PM site, ternary complexes were formed. This shows that the sequences of all three PH sites are capable of binding Pbx/Meis and PH *in vitro* in a direct and sequence-

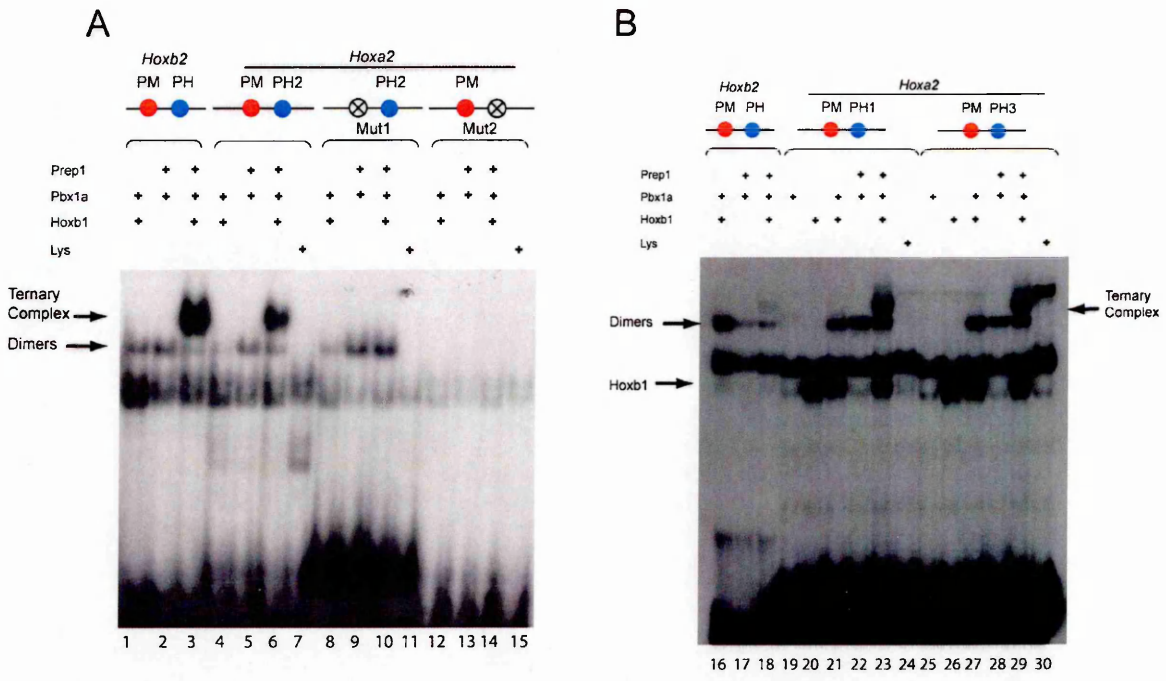


Figure 3.5. Electrophoretic mobility shift assays on the different sites of the highly conserved *Hoxa2* region.
The double stranded oligonucleotide was mixed with different combinations of Prep1, Pbx1a, and Hoxb1 protein (noted above the panels). In **A** and **B,** the first three lanes (in **A** lanes 1-3, in **B** lanes 16-18) represent the control experiment with the *Hoxb2* Pbx/Hox and Pbx/Meis binding sites (Ferretti et al., 2000). (**A**) In the following lanes (lanes 4-15), various forms of the *Hoxa2* Pbx/Meis and Pbx/Hox#2 sites have been analyzed. Mut1 and Mut2 are mutant forms of the oligonucleotide in the *Hoxa2* Pbx/Meis and PH sites, respectively. **(B)** In the next lanes (lanes 19-30), the PH #1 and #3 sites of *Hoxa2* were analyzed. EMSA performed by Dr. Ferretti.

specific manner; stable dimer complex formation of the first and third sites are dependent on the presence of a PM site, whereas the second PH site is not. Ternary complexes at all three PH sites are dependent on the PM site.

3.3.4 *In vivo* deletion analysis of the *Hoxa2* r4 enhancer elements in chick and mouse

To test the importance of these sites *in vivo,* I performed site-directed mutagenesis experiments and compared the enhancer activity of different mutant and wild type elements in either the chick embryo electroporation or transgenic mouse embryo assay system (Figure 3.6). For the chick experiments, I used the chicken intron region (construct #5) for the deletion

experiments, as this fragment was capable of generating a robust and specific r4 pattern of expression. For the transgenic mouse assay I used the comparable mouse intron fragment (construct #7). Ninety percent of the embryos electroporated with the control construct (#5) displayed r4 specific reporter staining (see Figure 3.6B). Deletion of any of the three PH sites affected the intensity and specificity of expression in r4 (constructs #10, 11 and 12 in Figure 3.6D); deletion of the second PH site (PH2) has the greatest effect (construct #11), reducing the percentage of embryos showing r4 specific expression to 22% (Figure 3.5C), whereas deleting the first (construct #10) or third PH (construct #12) reduced it to about 45% and 35%, respectively.

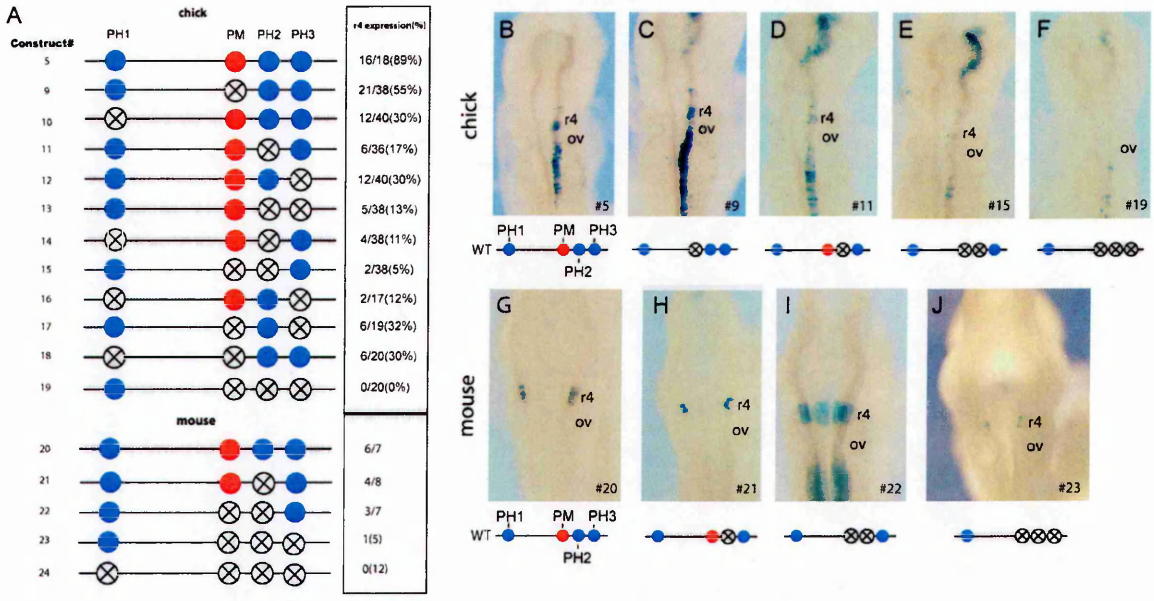


Figure 3.6. Analysis of PH and Pbx/Meis sites in chicken and mouse.
(A) Diagram shows mutation of each site or in various combinations in the chicken Hoxa2 intron region. Each construct is numbered on the left; the number of embryos electroporated with each construct and the frequency of r4 specific *lacZ* staining are presented on the right. (**B-E**) Representative electroporated embryos generated with variety of constructs (**B** #5, **C** #9, **D** #11, **E** #15, **F** #16). **(F-I)** Ventral view of representative transgenic mouse embryos showing reporter staining directed by various *Hoxa2* r4 enhancer constructs (**G** #19, **H** #20, **I** #21, **J** #22). The number of the construct used for the electroporation is shown in the left corner in each figure and in addition schematic diagrams of each construct is shown below each figure. r4, rhombomere 4; ov, otic vesicle.

Deletion of the Pbx/Meis site (construct #9) led to a reduction in the number of embryos displaying r4 specific expression to 63% (Figure 3.6D). Deleting the first Pbx/Hox or third PH site together (construct #16) also showed a reduction (11%) of r4 reporter activity, almost abolishing r4 enhancer activity. Deleting both the Prep/Meis and the second PH site (construct #15) led also to a great reduction of r4 enhancer activity (11%) (Figure 3.6E). De-

leting all three PH binding sites (construct #19) led to the complete loss of r4 enhancer activity. This implies that all PH binding sites contribute to r4 activity of *Hoxa2*, with the second PH site playing the most predominant role.

I also tested the contribution of each of the different elements for *Hoxa2* r4 enhancer activity in transgenic mouse embryos. I used the mouse *Hoxa2* intron as the basis for the mutation analysis (Figure 3.6G). Deleting the second PH site has only a minor effect in r4 enhancer activity, observing a reduction to 50% (Figure 3.6H). Deleting the Pbx/Meis in addition to the second PH site did not result in much more reduction of the r4 enhancer activity, which was different in the chick electroporation assay system (Figure 3.6I). Deleting the PH3 site, in addition to the PH2 and PM sites, led to a major reduction in the ability of the construct to direct r4 expression; the embryo showed only a few cells with positive reporter activity in r4 (Figure 3.6J). When all the elements are deleted no reporter activity was observed (construct #5).

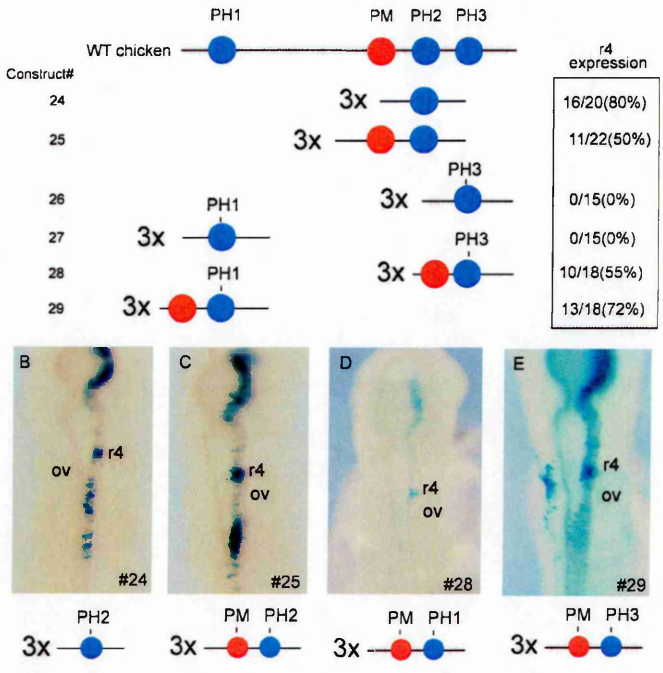


Figure 3.7. Analysis of different multimerized constructs in electroporated chick embryos.
(A) Diagram shows multimerized elements of each construct. On the right, the total numbers of embryos are noted and the frequency of r4 specific expression in numbers and percentage. **(B-E)** Representative electroporated embryos stained with *lacZ* (B #24, C #25, D #28, E #29).

3.3.5 Multimers of each PH element supports differential contributions to the *Hoxa2* r4 enhancer

To gain further insight into the role of each *Hoxa2* r4 element, I generated a series of constructs carrying oligonucleotides in which each of the identified elements were multimerized and linked to a *lacZ* reporter gene. Electroporation of the construct in which the PH2 site

has been multimerized, in either the presence or absence of the PM site (constructs #24 and #25), led to strong reporter expression in r4 in 50% and 80% of the electroporated embryos, respectively (Figures 3.7B,C); in contrast, embryos electroporated with a construct in which either the first or the third PH sites have been multimerized (constructs #26 and #27), showed no r4 reporter staining (data not shown). However, when the first or third PH is fused with the PM binding site, the activity is restored and the electroporated embryos showed r4 specific reporter staining (Figures 3.7D, E). This is consistent with the EMSA results, in which the ternary complex formation with the first and third PH sites depends upon the presence of the PM sites (Figure 3.4B).

3.3.6 Overexpressed Hoxb1 *trans*-activates the *Hoxa2* r4 enhancer *in vivo*

To further confirm that the *Hoxa2* r4 enhancer activity is Hoxb1-dependent, I overexpressed Hoxb1 throughout the hindbrain in chick embryos by electroporation (Figure 3.8). For this purpose, I used a CMV promoter/vector construct containing the human *Hoxb1* cDNA. This construct was co-electroporated with the *Hoxa2* r4 reporter construct (construct #5), and I observed in the majority of the embryos (11 out of 18 embryos) a *trans*-activation of the *lacZ* reporter staining in the hindbrain (Figure 3.8B). This observation is consistent with previous findings that *Hoxb1a* overexpression led to ectopic expression of *Hoxa2* in more anterior structures in zebrafish (Hunter and Prince, 2002).

3.4 Discussion

3.4.1 *Hoxa2* r4 regulation is integrated in the *Hoxb1* regulatory network in r4

In this study, I have shown that *Hoxa2* r4 expression is regulated directly by Hoxb1 and its co-factors Prep and Pbx. I used sequence and functional comparison among different species to prove that this is a conserved mechanism.

A global sequence alignment of the *Hoxa2/a3* intergenic and *Hoxa2* coding sequence region has allowed me to identify a highly conserved region in the *Hoxa2* intron that directs r4 specific expression. This control region contains three PH binding sites and one Prep/Meis site, which are able to form ternary complexes with Hoxb1, Pbx and Prep/Pbx proteins.

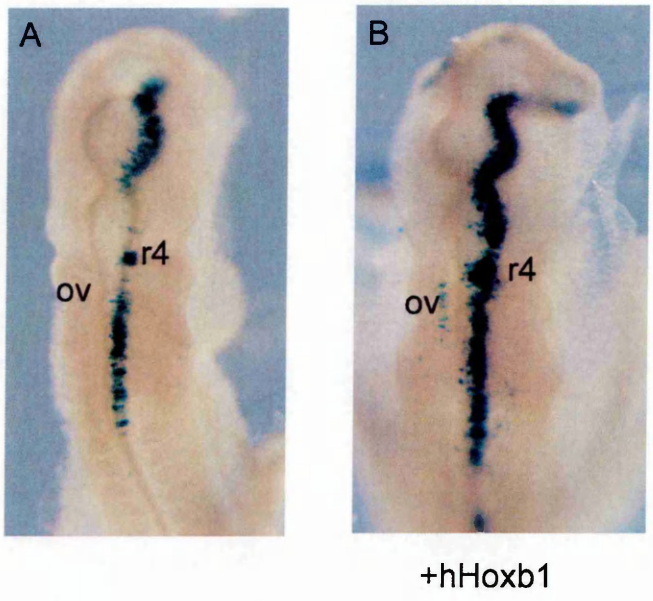


Figure 3.8. The r4 enhancer of *Hoxa2* is *trans*-activated *in vivo* by Hoxb1.
Control embryo electroporated with the *Hoxa2* r4 reporter construct (construct #5), showing r4 restricted expression in the hindbrain. **(B)** Embryo which has been electroporated with the *Hoxa2* r4 enhancer and expression construct containing the human *Hoxb1* gene driven by a CMV promoter.

I provide evidence to support that *Hoxa2* r4 expression is directly regulated by Hoxb1, which involves Pbx and Prep/Meis proteins as cofactors. First, the *Hoxa2* intron contains blocks of conserved sequences, which contain PH and Prep/Meis binding sites. Second, this conserved region is able to direct r4 specific reporter expression *in vivo*. Third, deletion analyses in mouse and chick revealed that all binding sites are essential for robust r4 reporter expression. Finally, overexpression of Hoxb1 in the chick hindbrain *trans*-activates the *Hoxa2* r4 enhancer in the hindbrain. Altogether, these results strongly suggest that the *Hoxa2* r4 enhancer is located in the *Hoxa2* intron and that r4 expression of *Hoxa2* is regulated directly by Hoxb1.

Hoxb1 plays a major role in r4 patterning and establishment of the r4 identity (Figure Studer et al., 1996; Studer et al., 1998). Its expression is initiated by retinoids and maintained by a conserved autoregulatory control element. In addition, through a direct cross-regulatory mechanism, Hoxa1 and Hoxb2 participate in maintaining *Hoxb1* r4 restricted expression and, therefore, the identity of r4. Hoxb1 has been shown to directly regulate *Hoxb2* expression in r4 (Maconochie et al., 1997). This means that Hoxb1 is an important component in the genetic

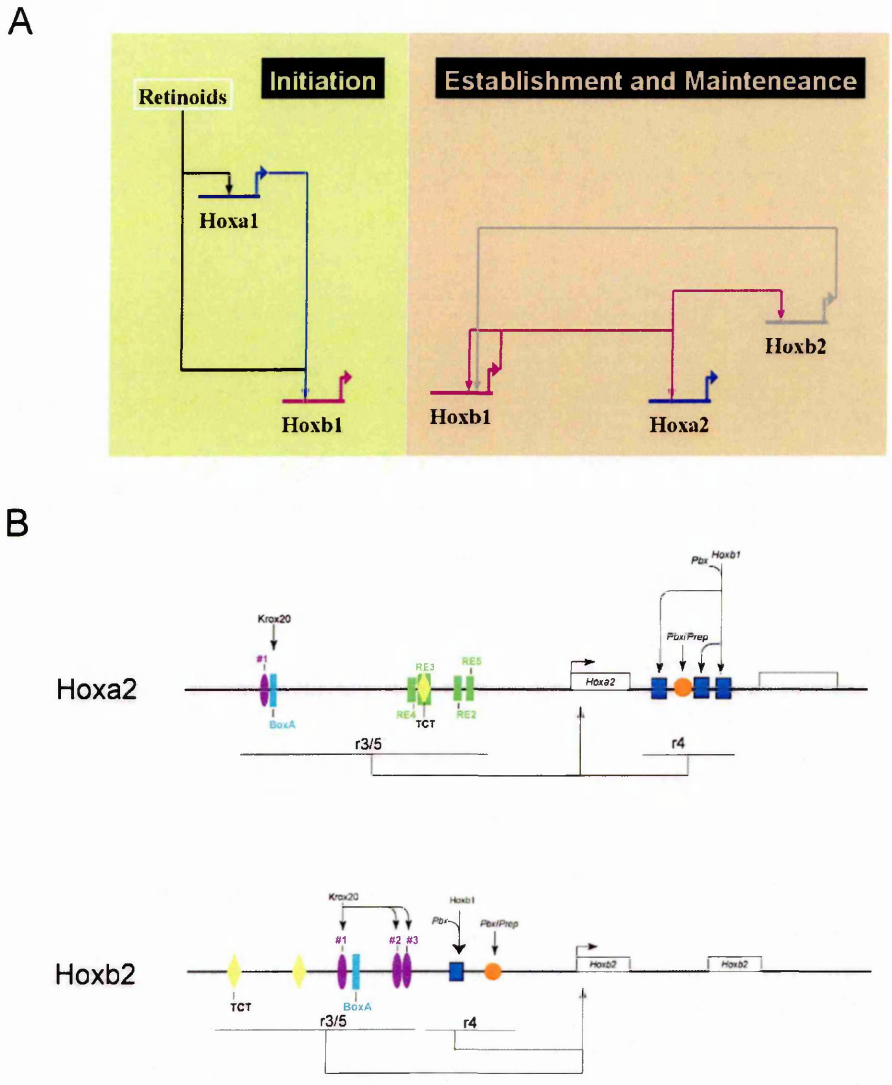


Figure 3.9. The rhombomere 4 network and comparison of the regulatory modules of the paralogous members *Hoxa2* and *Hoxb2*.
(A) *Hoxa2* is a component of the *Hoxb1* regulatory network in rhombomere 4. Dotted line illustrates that the interaction between the genes have been not shown to be direct. Arrowhead at the end of the line represents activation, whereas perpendicular line at the end of the line illustrates repression between the connected genes. **(B)** Comparison of the *Hoxa2* and *Hoxb2* regulatory modules for r3/5 and r4 expression. Note that in both cases the r3/5 regulatory elements are located in the intergenic region, whereas the r4 elements directing r4 expression of *Hoxa2* are located in the intron and the elements for *Hoxb2* expression adjacent to the r3/5 module in the intergenic region.

regulatory network in r4, and this study allows us to further refine the r4 regulatory network by adding *Hoxa2* has another target gene of Hoxb1 (Figure 3.9A).

Hoxa2 is expressed more strongly in r4 of chick embryos than in mouse embryos (Figure 3.1). This suggests that there are subtle differences in the *Hoxa2* r4 elements between chick and mouse, which is reflected in the inter-species experiments (Figures 3.2E, F). However, the conservation of the *Hoxa2* r4 elements is remarkably high among mammals and chick. Only the PH2 shows subtle sequence divergence between chick and mammals (Figure

3.3B). This suggests that flanking regions have additional input for r4 enhancer activity which account for species-specific expression. Also, the deletion analysis suggests that there are slight differences in the contribution of activity of each of the *Hoxa2* r4 enhancer elements between mouse and chick (Figure 3.6). In the mouse *Hoxa2* r4 enhancer, the PH1 and 3 sites play an important role; deleting both of them led to a complete absence of enhancer activity, whereas the PH2 and PM sites plays only minor roles in this enhancer. In chick, the PH2 and PM sites plays the major role in directing r4 expression. These functional differences between the mouse and chick r4 *Hoxa2* enhancers can be due to the assay system or, more likely, to species specific differences in the contribution of the elements.

3.4.2 Conservation and diversity of the regulatory elements of *Hox* group 2

Both paralogous members *Hoxa2* and *Hoxb2* are expressed in r3 up to r5. *Hoxa2* exhibits additional expression in r2. The expression in r3 and r5 of both genes is regulated by Krox20-dependent enhancers located in their upstream intergenic regions(Nonchev et al., 1996b; Sham et al., 1993). However, the *Hoxb2* r4 enhancer is located adjacent to the r3/5 module and, in this study, I showed that the *Hoxa2* r4 enhancer is located in the intron and that both modules are integrated in the *Hoxb1* regulatory network (Figure 3.9B).

This raises the interesting question whether the regulatory regions evolved prior the duplication of the *Hox* genes, or whether they evolved after the duplication event. A third scenario could be possible, in which either of the r4 elements of *Hoxa2* or *Hoxb2* changed its location after they were duplicated.

The presence of conserved Krox20 sites in both *Hoxa2* and *Hoxb2* intergenic regions supports the first and third hypotheses, although the diverged location of the Hoxb1-dependent r4 regulatory region in *Hoxa2* (intron) and *Hoxb2* (intergenic region) supports the latter statement. It is possible that the Krox20-dependend r3/5 modules were present prior the duplication of the *Hox* clusters and that the r4 enhancers evolved independently using the same upstream

factors. This view is further supported by the fact that the numbers of PH sites differ between the two modules: there are three PH sites in the *Hoxa2* r4 module and one in the *Hoxb2* r4 module.

3.4.3 Hox, Prep, and Pbx binding sites are present in various control elements and are able to direct tissue specific expression

The expression of several *Hox* genes has been shown to be dependent on PH and Prep/Meis regulatory elements. Although these PH sites are very similar to each other, they direct expression in different domains; for example, the PH sites of *Hoxb1*, *Hoxb2,* and *Hoxa2* direct expression in r4, whereas the *Hoxa3* PH site directs expression in r5/6. This suggests that variation within the PH binding site or in the intermediate flanking sequence can have an effect in their output activity *in vivo*. Further biochemical and transgenic analysis must be performed to determine the nature of these differences.

Chapter 4

Regulation of *Hoxa2* expression in Rhombomere 2

4.1.1 Introduction

Hox genes play an important role in establishing the regional identity throughout the anteroposterior (AP) axis of embryo (Krumlauf, 1994; Lumsden and Krumlauf, 1996). This is achieved by the *Hox*-code; the combinatorial distribution of different Hox proteins provides cells with molecular information about their AP position (Hunt et al., 1991). This has been shown extensively in the hindbrain, which is segmented into lineage-restricted compartments, termed rhombomeres (r). Several *Hox* genes are expressed in the hindbrain with overlapping expression patterns. For example, *Hoxb1, Hoxb2* and *Hoxa2* are expressed in r4, whereas only *Hoxa2* and *Hoxb2* are expressed in r3; in r2 only *Hoxa2* transcripts are present and r1 is considered to be a '*Hox*-free' domain. Disruption of this code has been shown to change the identity of the rhombomere. The targeted inactivation of the *Hoxa2* gene, which is the only *Hox* gene expressed in r2, leads to a change in the expression patterns of various genes and morphological abnormalities in r2, to such an extent, that it results in a partial transformation to an r1 identity (Gavalas et al., 1997). The basis for this conclusion is that *Sax-1* expression, which is upregulated in the ventral portions of the r1 in WT embryos, is extended posteriorly into the r2 domain in mutant animals. Further it has been shown that *Sek-1,* normally expressed in r2, is absent in *Hoxa2* deficient mice (Gavalas et al., 1997).

In WT embryos all of the r2/3 motor axons exit through the r2 and innervate the first branchial arch (BA). In the *Hoxa2* mutant, some of r2/3 motor neurons originate, instead, from r4 into the transformed second BA. These mutant phenotypes occur in addition to the loss of the r1/2 border. However, the r2 trigeminal exit point is not changed in *Hoxa2* mutants, suggesting that the transformation is not complete.

In order to understand the molecular and morphological changes caused by the disturbance of the *Hox* code, elucidating the regulatory mechanisms of individual *Hox* gene expres-

sion is essential. It has been shown that r3/5 expression of *Hoxa2* is regulated by a complex mechanism which involves the zinc-finger transcription factor, Krox20. Further, it has been shown that several elements are responsible for *Hoxa2* expression in NCC and that the transcription factor AP2 is directly involved in this process. In the previous Chapter (Chapter 3) of my thesis, I have demonstrated that the r4 enhancer is located in the *Hoxa2* intron and that its expression is dependent on Hoxb1 and its cofactors Prep/Pbx, but the mechanisms which underlie the regulation of *Hoxa2* expression in r2 are unknown.

In this study, I show that the *Hoxa2 cis*-regulatory elements necessary for r2 expression are located in the second exon of *Hoxa2* using chick electroporation and mouse transgenic assays. This highly conserved enhancer consists of several components, which can be subdivided into maintenance and initiation elements. I present data which show that the r2 enhancer specific sequences are conserved beyond the level required for amino acid coding requirements. Finally I show that the expression in r2 is not regulated by an auto-regulatory mechanism.

4.2 Results

4.2.1 Mapping a conserved r2 enhancer in Exon2 of *Hoxa2*

It has previously been shown that the intergenic region of *Hoxa2/a3* contains the r3/5 and NCC enhancer (Nonchev et al., 1996b). In the previous Chapter (Chapter 3) I showed that the intron region contains the r4 enhancer. In order to identify the r2 regulatory elements, I analyzed the more 3' region of the *Hoxa2* gene locus. This region has been shown in initial studies to contain the r2 regulatory region (Frasch et al., 1995). I cloned and sequenced the chicken 3' downstream region, linked it to a *lacZ* reporter gene, and used chick electroporation and mouse transgenic analysis as assay systems for testing r2 regulatory activity (Figure 4.1). A construct containing a 1.9kb *MluI-BamHI* fragment (#1), which includes the second half of *Hoxa2* exon 2 and additional 3' downstream sequence, directs strong reporter expression in r2 in chick embryos (n=20) (Figure 4.1B). After more precisely defining the region containing

the r2 enhancer (constructs #2-4), I demonstrated that the r2 regulatory region is located in the second exon of *Hoxa2* (Figures 4.1A, C). Constructs #2 to #4 gave constitently r2 specific staining in the chick (Figures 4.1B, C and data not shown). Additionally, construct #2 showed strong r2 expression in transgenic mice (Figure 4.1D). Further examination of this region showed that construct #5, but not #6, was directing reporter expression in r2 in the electroporated chick embryos and in transgenic mice (Figures 4.1E, F). Hence, the *Hoxa2* r2 module is located in the 3' region of the second exon of *Hoxa2*.

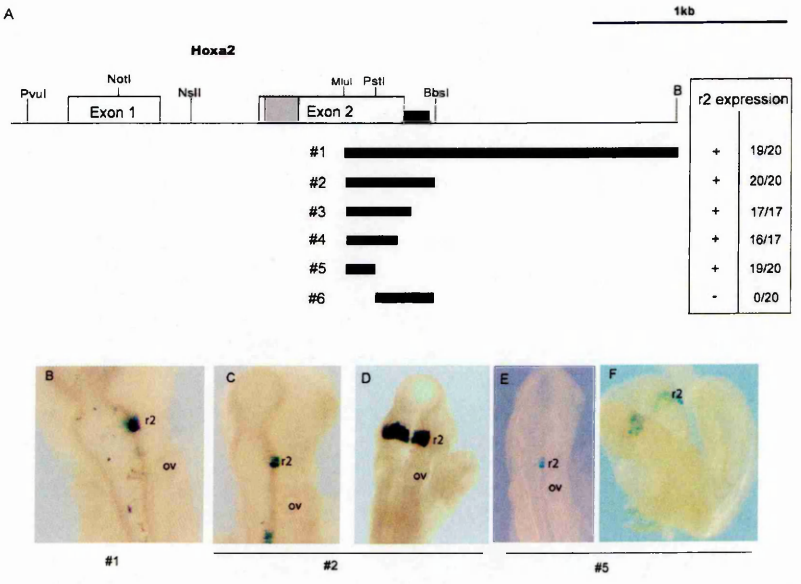


Figure 4.1. Identifying the *Hoxa2* module directing r2 expression. **A.** Schematic diagram showing the relative location of the construct used for identifying the r2 enhancer. The grey box represents the location of the homeodomain and the black shows the position of the 3'UTR. **(B-F)** Representative electroporated chick **(B, C, E)** and transgenic mouse **(D, F)** embryos used for identifying the r2 regulatory module.

I also cloned the regions equivalent to construct #2 from other species, including mouse, zebrafish, dog, and frog *(Xenopus tropicalis)* and show that they also direct reporter staining in r2 (Figures 4.2A-D) suggesting that the r2 module is highly conserved among these different species.

4.2.2 Multiple *cis*-acting elements are required for regulating *Hoxa2* expression in r2

I used various constructs to further identify the precise location of the r2 module of *Hoxa2*. I designed this experiment by making different constructs, deleting progressively larger regions from both either 5' or 3' ends (Figure 4.3). Constructs #7 through #11 were deletions of regions starting at the 5' end (210bp, 190bp, 160bp, 140bp, and 100bp), and

constructs #12 through #14 begin at the 5' region with 3' deletions (200bp, 150bp and 110bp). I linked these constructs to the *lacZ* reporter gene and used chicken electroporation as an assay. Constructs #7, #8, #9, #10 and #12 were able to direct r2 restricted reporter expression. However, the efficiency progressively decreased from construct #7 to #8 (from 89% to 50%) and construct #8 to #9 (from 50% to 35%). A region located between 100bp and 200bp is sufficient to initiate r2 expression (see blue box), however there are other elements maintaining or enhancing r2 expression, located in the region between 50bp and 100bp, which I termed maintenance elements (see red box) (Figure 4.3), because they affect the level of expression.

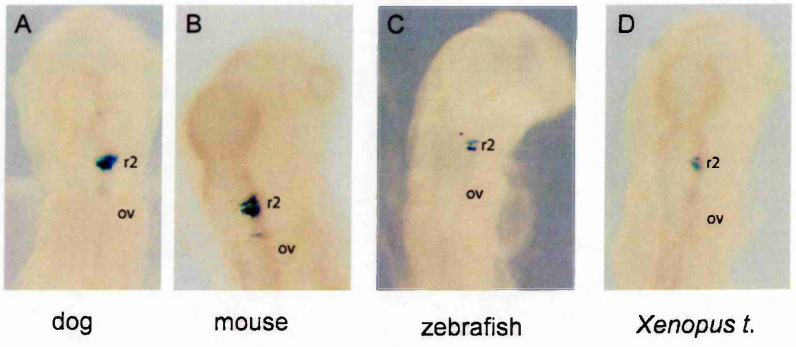


Figure 4.2. Analyzing the r2 enhancer of different species.
(A-D) Electroporated embryos tested with the *Hoxa2* r2 enhancers from different species including dog, mouse, zebrafish and frog (*Xenopus tropicalis*).

To screen for the elements involved in initiating and maintaining r2 expression, I performed deletion experiments in the 170bp region, spanning the region from 50bp to 220bp of construct #2 (Figure 4.4). For the first 50 bp I performed five 10bp (#15, #16, #19, #20 and #23) deletions, and for the following 120bp, I performed six 20bp deletions (#26, #27, #28, #31, #32 and #35). The deletions of constructs #16(Δ2), #20(Δ4), #23(Δ5), #29 (Δ8) and #32 (Δ10) led to a reduction in r2 enhancer activity. Whereas the other deletions led to no remarkable reduction of r2 reporter activity, showing roughly 90% or greater expression.

In order to map the precise location of the r2 elements, I performed additional 5bp and 10bp deletions in the regions identified as having an effect on r2 enhancer activity (Figure 4.4). For each of the three 10bp regions (Δ2, Δ4, Δ5), I performed two 5bp deletions (Δ2.1, Δ2.2, Δ4.1, Δ4.2, Δ5.1, Δ5.2), and for each of the two 20 bp regions (Δ8 and Δ10), I deleted two 10 bp regions (Δ8.1, Δ8.2, Δ10.1, Δ10.2). Of the first 6 deletions, Δ2.1, Δ4.1, Δ4.2 and Δ5.2 resulted in the loss of expression in r2. Deleting the first 10bp of region of Δ8 had no ef-

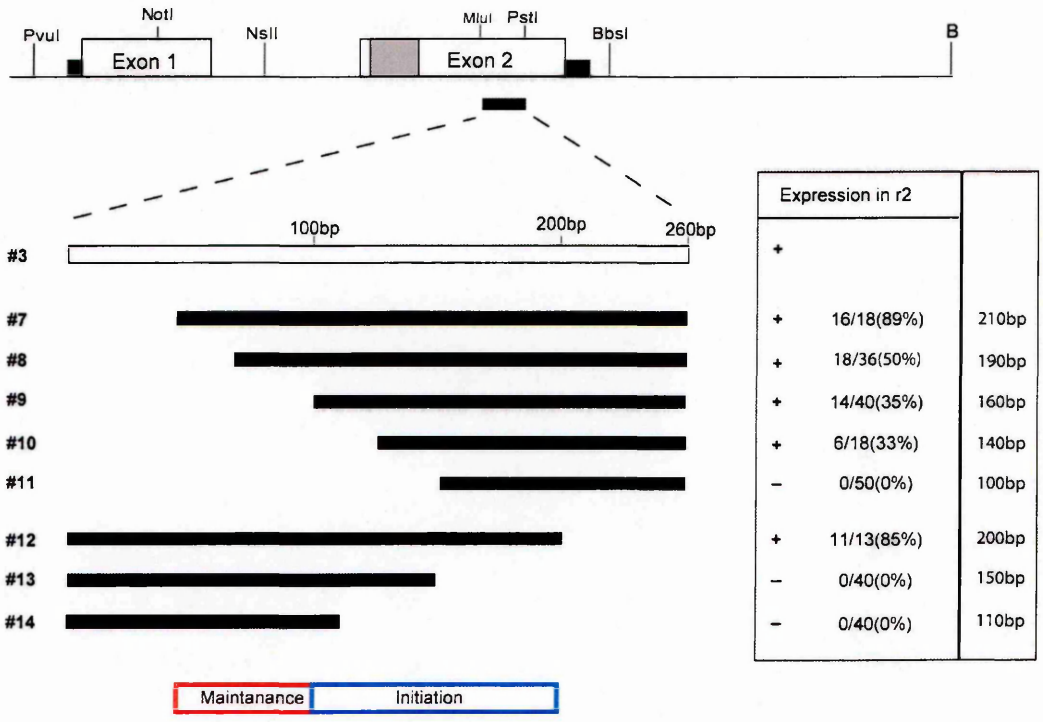


Figure 4.3. Constructs for electroporation analysis in mapping the r2 module of *Hoxa2*.
At the top is a schematic diagram of the *Hoxa2* locus and below are the constructs listed that were linked to a *lacZ* reporter gene for regulatory analysis. Construct #3 is used as a template for all sequential constructs. Construct numbers are noted on the left. The efficiency of expression in r2 and the size of each construct are listed on the right. Below the constructs, boxes illustrate the location of the maintenance (red) and initiation (blue) regions.

fect on r2 expression, whereas deletion in the second 10bp led to a reduction of around 60%. Deleting the first 10bp of region Δ10 showed great reduction (reduction to 35%) of r2 enhancer activity, whereas deleting the second 10bp did not have any effect on r2 enhancer activity. Through this analysis, I identified at least three elements embedded in the maintainance region (red) controlling levels of expression and two elements in the initiation region controling specificity (blue), which mediate r2 expression of *Hoxa2*.

4.2.3 Alignment of the region containing the r2 elements

Having identified specific regions required for expression of *Hoxa2* in r2, I performed an alignment comparing the sequences from various species (Figure 4.5). The three red boxes encompass the elements identified by the deletion experiments in the maintenance region, whereas the two blue boxes contained the elements initiating r2 expression. It is possible to define the consensus sequence for each element; for this purpose I did not include the sequence of *Hoxa2(a)* medaka and fugu, since it is been shown that the expression of fugu *Hoxa2(a)* is very low in r2 (see also chapter 5 and Amores et al., 2004). I termed the first three elements RTE1-3 (for **r**hombomere **t**wo **e**lement 1-3). The RTE1 element shows a very high degree of conservation; its consensus sequence can be defined as VCCAR. The second RTE motif shows less conservation. In most species there is a conserved CCTTTANC core, which divergences in the fish. The third RTE motif is extremely highly

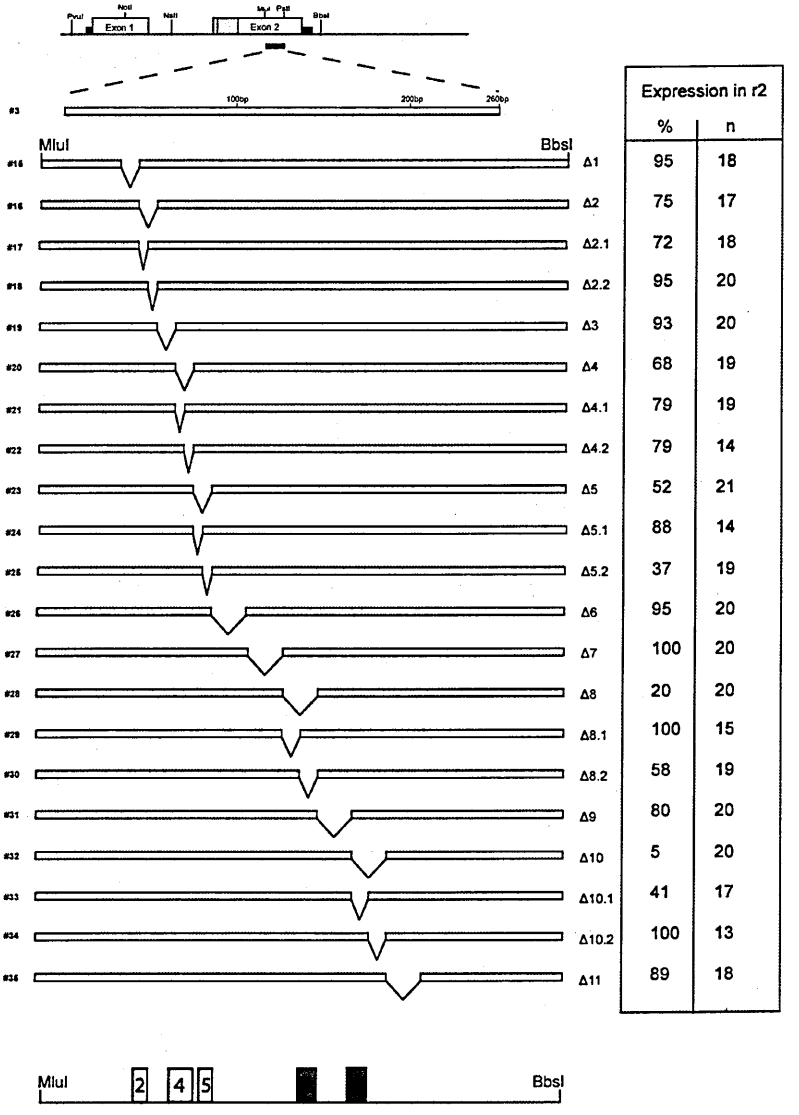


Figure 4.4. Deletion analysis of the r2 module for identifying the r2 elements of *Hoxa2*.
At the top is a schematic diagram of the *Hoxa2* locus, below are the constructs listed, which are used for mapping the r2 elements. The deletions were performed in the area spanning construct #15. Not that the deletions were performed in the fragment used for construct #2. On the left the construct numbers are listed. On the right the deletion numbers and the efficiency are noted. On the bottom of the figure the elements are shown which are important for r2 enhancer activity, the colors of the elements illustrate whether they belong to the maintenance (red) or initiation (blue) region.

conserved, with the consensus sequence of GAVAA. The blue boxes contain very similar and highly conserved ACAAT motifs (see orange boxes) and will be defined as "ACAAT motifs" hereafter. These data suggest that the two ACAAT motifs may mediate the initiation of *Hoxa2* expression in r2. I therefore analyzed the *in vivo* activity of these individual motifs (Figure 4.6). I generated a series of constructs carrying multimerized double-stranded oligonucleotides (4 copies), including the ACCAT motifs linked to a *lacZ* reporter and assayed using chick electroporation. The first multimerized construct (#36) contains four copies of the region deleted in Δ8. The majority of embryos electroporated with this construct show strong expression in r2, although weak reporter staining was also observed in adjacent rhombomeres.

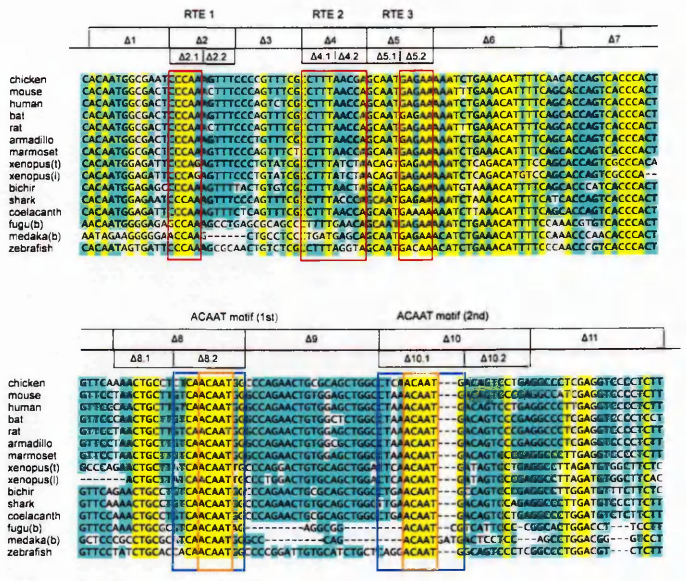


Figure 4.5. Sequence alignment of the *Hoxa2* r2 module. The boxes within the sequence show the important elements for r2 enhancer activity. On the top of the sequence, the numbers of the deletions (Δ1 to 11) and the names of the elements (RTE1-3 and ACAAT motifs 1-2) are positioned.

Next, I tested the region deleted in Δ10 (#37), containing the second ACAAT motif. Here, I also observed strong r2 reporter expression in the hindbrain (Figure 4.6C). In order to define the minimal sequence necessary for directing r2 expression, I used shorter sequences encompassing the second ACAAT motif (#38 and #39). I determined that multimerization of an 11bp region encompassing the ACAAT motif is able to direct *lacZ* expression in r2 (Figure 4.6D). Surprisingly, by just multimerizing the ACAAT sequence alone, 40% of the embryos tested showed r2 expression (Figure 4.6E). This suggests that this motif is sufficient for initiating r2 expression.

4.2.4 Changing the sequence of the *Hoxa2* r2 enhancer while maintaining the Hoxa2 encoded amino acid sequence

A surprising finding of this study is that the *Hoxa2* r2 elements are embedded in the second exon of *Hoxa2*. This sequence is highly conserved as expected for a coding exon (Figure 4.5). Drift of this sequence could be constrained by both the requirement to maintain the amino acid sequence of the protein and the requirement to maintain regulatory elements directing r2 expression. In order to not compromise the function of the encoded Hoxa2 protein, I altered the proposed r2 elements using bases which would maintain the encoded amino acids (for example, in the third base, the so-called wobble position). Side-directed mutagenesis in the ACAAT motifs had the strongest effect, whereas changes in the third bases of a codon of the RTE motifs had small or negligible effects (Figure 4.7). I also changed the second nucleotide of several codons embedded in the ACAAT motifs. I subdivided the effect in three categories: no change of r2 activity (blue in Figure 4.7, 81-100%), medium effect in which the number of the tested embryos showing r2 specific reporter staining was reduced to 51 to 80% (yellow in Figure 4.7) and great effect in which only 17 to 50% of the tested embryos show r2 reporter expression(red in Figure 4.7).

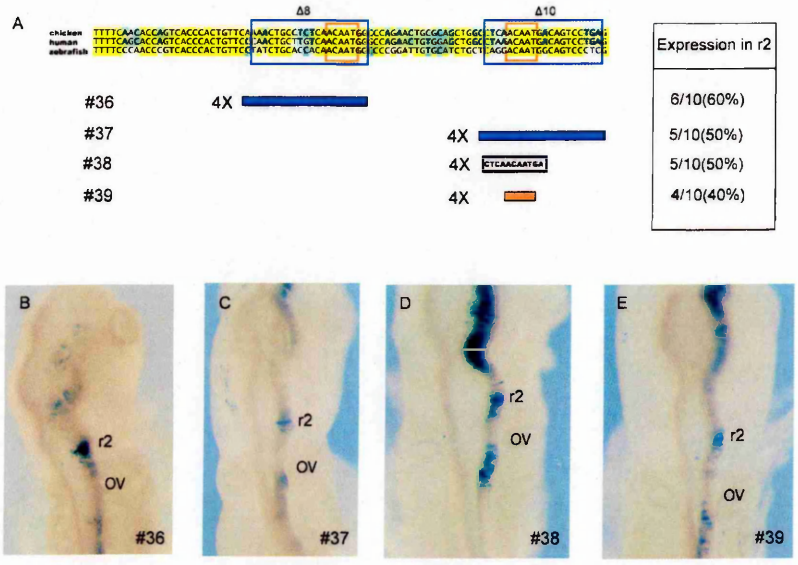


Figure 4.6. Analysis of regulatory activity of the individual ACCAT motifs.
(A) On the top a sequence alignment is displayed used for analysis of the ACCAT motifs. The blue box encompasses the sequenced defined by the deletion experiments (Δ8 and 10), the red box shows the location of the ACCAT motifs. The constructs listed below are used for this experiment. Four copies of an oligonucleotide for the ACAAT region in various lengths are fused to the reporter construct. The blue constructs contain the region of defined by the deletion experiments (20bp) (construct #36 and 37), whereas the black construct contains a smaller region (11bp) (construct #38). The red construct contains only the ACAAT motif (5bp) (construct #39). On the left the construct numbers are displayed and on the right the efficiency is listed. **(B-E)** Embryos electroporated with the above construct stained for reporter activity.

The first ACAAT motif of chick is located in a region which contains three codons:

TCA for the serine, ACA for threonine and ATG for methionine (Figure 4.7). For the first and second codon, I was able to change the third nucleotide without altering the encoded amino acid (serine). Changing the TCA to TCG had no effect on the r2 enhancer activity; this is consistent with the second ACAAT motifs of some fishes (fugu, zebrafish and medaka), which have a G at this position. But changing this codon from TCC or TCT reduces its activity in a medium range. The second codon (ACA) can be changed to ACC, ACG, and ACT without changing the encoded amino acid. The first two changes (to ACC and ACG) had a profound effect on its activity, whereas the change to ACT had only a medium effect.

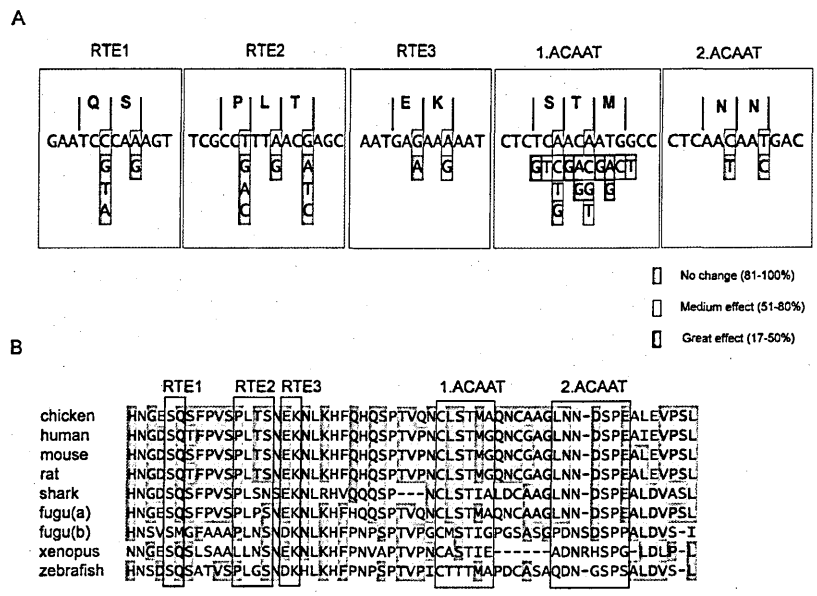


Figure 4.7. Analyzing the sequence of the r2 elements by changing the third base of the codon.
(A) Schematic diagram of the changes introduced into the *Hoxa2* r2 elements. The letters on the top of each panel are the amino acids encoded by the codon sequences listed below. Below are the single base pair changes of the third base of the codon. The colors illustrate the effect of the changes performed in each of the elements (blue: no change; yellow: 51 to 80% of all electroporated embryos show r2 restricted expression; red: 17 to 50 % of all electroporated embryos display reporter staining in r2). **(B)** Sequence alignments of the protein encoded by the r2 module.

The third codon in the first motif is methionine and no nucleotide changes can be performed to retain the encoded amino acid. I therefore asked what happens when I changed the codon sequence along with the encoded amino acid. Changing the ATG to ATC (isoleucine) has a medium effect, whereas changing the codon sequence from ATG to GTG (valine) had also a medium effect on the ability of the element to direct *lacZ* expression in r2.

I performed the same series of experiments for the second ACAAT motif. Here the alternative codon sequences were limited. The first codon, AAC, (asparagine) in this region, can be changed to AAT without affecting the encoded amino acid. The second codon, AAT, embedded in the second ACAAT motif also encodes for asparagine. It can also be changed to

AAC without affecting the encoded amino acid. Both changes had no effect on r2 enhancer activity.

As mentioned before, several fishes show the consensus sequence **G**ACAAT instead of **A**ACAAT for the second ACAAT motif. Therefore their first codon encodes aspartic acid, instead of asparagine. This means that to keep the first two nucleotides of the ACAAT motif conserved, both codons, AAC for asparagines and GAC for aspartic acid, are tolerated at that position for creating a functional *Hoxa2* protein.

4.2.5 *Hoxa2* expression in r2 is not maintained by an auto-regulatory mechanism

To gain further insights into *Hoxa2* r2 expression in the hindbrain, I asked whether *Hoxa2* is required to maintain its own expression. This regulatory mechanism has been observed in several *Hox* gene members, including *Hoxb4*, *Hoxb1* and *Hoxa3* (Gould et al., 1997; Manzanares et al., 2001; Pöpperl et al., 1995). I therefore crossed the *Hoxa2* mutant into a stable transgenic line carrying the mouse r2 enhancer linked to a human alkaline phospatase reporter gene. I observed no differences in reporter staining between the WT and *Hoxa2* mutant, which shows that *Hoxa2* expression is not maintained by an auto-regulatory mechanism in r2 (Figure 4.8).

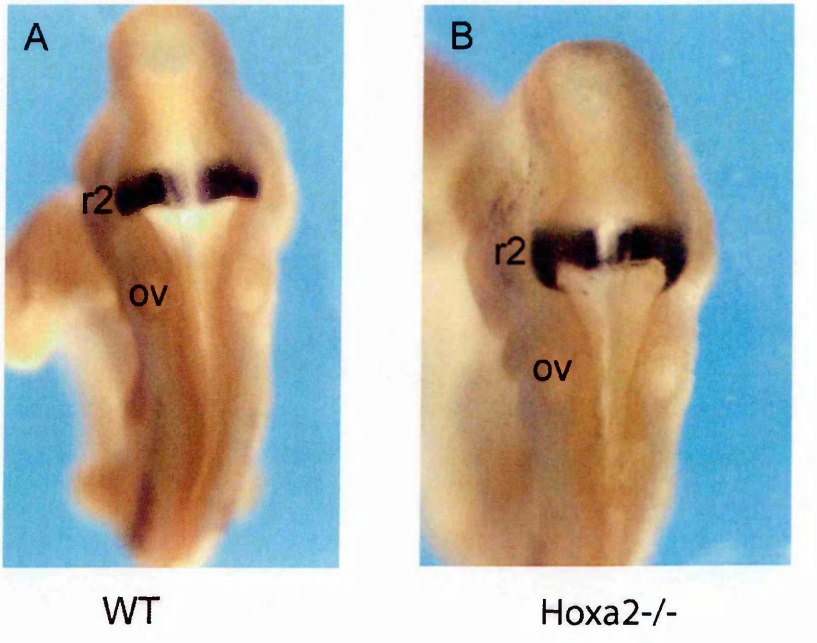


Figure 4.8. Analyzing the r2 enhancer in the mouse *Hoxa2* mutant. **(A)** Control transgenic mouse carrying the mouse *Hoxa2* r2 enhancer linked to a human alkaline reporter construct. **(B)** *Hoxa2* mutant carrying the same transgene as in **A**.

4.3 Discussion

4.3.1 The *Hoxa2* r2 enhancer elements are located in exon2

Different mechanisms have been described for regulation of *Hox* gene expression during development, including para-, cross- and autoregulation (Gould et al., 1998; Nonchev et al., 1996b; Pöpperl et al., 1995). In most cases the *cis*-regulatory elements directing *Hox* expression are located in the upstream intergenic region (Haerry and Gehring, 1996; Manzanares et al., 1997; Manzanares et al., 2001; Nonchev et al., 1996b). *Hox* gene enhancer elements have also been identified in the intron region for *Hoxa2* and *Hox* genes of the fourth paralogous group (see last chapter and Aparicio et al., 1995; Brend et al., 2003; Haerry and Gehring, 1996). In this study I demonstrate a novel position for an element which regulates *Hox* gene expression; I showed that the r2 *Hoxa2* enhancer is located in the second exon of *Hoxa2*. The location of an enhancer in the coding sequence has been recently reported in the Bcl-2 gene(Lang et al., 2005), but this regulatory mechanism is the first of its type described for *Hox* genes. This is important, since *Hoxa2* is the only *Hox* gene expressed in r2 and its expression has a crucial function in establishing the identity of this rhombomere(Gavalas et al., 1997).

This mechanism is highly conserved among different species. I have shown that the region containing the r2 enhancer of various species is able to direct reporter expression in electroporated chick embryos. I precisely located various elements important for r2 regulation and subdivided them into initiation and maintenance elements which are necessary to establish the r2 expression. The initiation elements are highly conserved, encompassing two ACAAT elements. Reporter constructs containing multimerized copies of this element were able to direct expression in r2. Further I performed changes in the r2 elements without changing the encoded amino acid. The results suggested that the bias codon usage is a result of the necessary to maintain functional r2 elements. Finally, I show that r2 expression is not regulated by auto-

regulatory mechanisms.

I show that the ACAAT motifs are able to initiate r2 expression, and the RTEs appear to maintain expression. By testing the multimerized ACAAT motifs in chick electroporation (Figure 4.6), there was strong expression in r2, but also additional expression in adjacent rhombomeres. Therefore an additional functional role of the RTE elements could be to restrict and specify expression to r2.

Targeted insertion of a *neo* cassette in the 3' untranslated region of *Hoxa2,* creates a hypomorphic allele and results in homozygous lethality and changes in gene expression patterns in the hindbrain (Ren et al., 2002). These authors showed that the insertion of the *neo* cassette led to downregulation of *Hoxa2* expression in r2 and the removal of the *neo* cassette by Flp-mediated excision, resulted in animals with no obvious defects (Ren et al., 2002). My finding that the *Hoxa2* r2 elements are located in the 3' region of the *Hoxa2* coding sequence, may explain the r2 specific down-regulation in these mutants to some extend. It is possible that the phosphoglycerokinase (PGK) influences the r2 elements, since it has been inserted only 462 bp away from the most 3' r2 element (2nd ACAAT). Another possibility would be that the *neo* cassette promoter is in a position to compete with the endogenous *Hoxa2* promoter for the r2 enhancer element.

4.3.2 Evolutionary implications of the *Hoxa2* r2 regulatory region

In this study, all analyzed *Hoxa2* r2 regulatory sequences were from gnathostomes, jawed vertebrates. In lamprey, a jawless (agnathan) vertebrate, it has been shown that the *Hoxa2* orthologous gene, *Hox2*, also displays the most anterior expression compared to the expression of other lamprey *Hox* genes (Takio et al., 2004). Analysis of the 3' coding sequence of *Hox2* reveals the presence of only one ACAAT motif, which fulfills the consensus sequence VRACAATD (data not shown). I was unable to identify any RTE motifs (data not shown). Therefore, it is likely that the expression of the anterior domains of *Hox2* is regulated in a different manner or/and that one ACAAT motif is sufficient to regulate the anterior ex-

pression of *Hox2*.

It is interesting to speculate how the region containing the r2 elements evolved, whether the bias lies toward creating optimal r2 regulatory elements or toward creating an appropriate Hoxa2 product. My site-directed changes in both ACAAT motifs may give some hints in this regard. I showed that changes in the ACAAT sequence can affect r2 enhancer activity, although the encoded amino acid sequence is not changed.

It has been noted that different species have a usage bias for certain codons (codon usage). I therefore checked the frequency of usage of codons embedded in the first ACAAT motif. The codon usage database was used (http://www.kazusa.or.jp/codon/). The first TCA codon is used less often (14.7%) than TCT (18.1%) or TCC (20.1%) in chicken. Changes of the sequence in this codon to TCT or TCC reduced the r2 activity (Figure 4.7). This suggests that the biased usage is due to the need to maintain proper functioning of the enhancer.

The second codon (ACA) in the first ACAAT motif, is the second most used codon in chicken (30.1% compared to ACT: 24.7%, ACC: 30.8%, ACG: 14.4%). Changes of the sequence from ACA to ACC or ACG resulted in a strong reduction of r2 enhancer activity. Therefore the occurrence of this codon could be due to using the most common codon, but also using the appropriate nucleotide for the r2 enhancer.

In summary, most changes, in which the sequence of the ACAAT motifs is changed and the encoded amino acid sequence retained, have an effect on r2 enhancer activity. Based upon these observations it appears that most codons in the two ACAAT regions are dictated by the need to maintain a functioning ACAAT motif, and the encoded amino acid sequence has just a secondary role. This view is also supported by the fact that the regions containing the two extremely highly conserved ACAAT motifs encode different amino acids due to their being used in a different codon frame.

4.3.3 Several *trans*-acting factors are involved in the Regulation of *Hoxa2* in r2

The ACAAT motif encompasses the consensus binding site for several *Sox* genes (Murakami et al., 1999). I tested the ability of several *Sox* genes, including *Sox2, Sox6, Sox7, Sox13,* and *Sox17,* to bind ACAAT elements in a luciferase reporter assay system. Two Sox gene members, Sox7 and Sox17, were able to activate reporter expression *in vitro* (data not shown), but no *Sox* gene have been described with restricted r2 expression to date.But, it may be possible that an uncharacterized Sox family member is involved in r2 regulation.

There are several genes expressed in r2, for example *Gbx2, Follistatin* and *Msx2*. But only a few genes have been show to be restricted to r2, e.g. *Cyp26c1*. This raises the question as to how expression in r2 is regulated, and which transcription factors mediate this restricted expression. My study suggests that several *trans*-acting factors are involved in regulation of r2 expression of *Hoxa2*, as I identified several different elements necessary for proper r2 regulation. The ACAAT motifs are necessary to initiate r2 expression. These two motifs are extremely similar, and it is likely that similar transcription factors are binding to these two elements, possibly as dimers. This initiated expression is most likely maintained and specified by different transcription factors binding to the RTE elements.

Other *Hox* genes maintain their expression by cross-, para- and auto-regulatory mechanisms (Gould et al., 1997; Manzanares et al., 2001; Pöpperl et al., 1995). I can rule out cross- and para-regulation of *Hoxa2*, because *Hoxa2* is the only *Hox* gene expressed in r2. I further show that auto-regulation is not involved in maintaining r2 expression, because the *Hoxa2* enhancer is still active in the *Hoxa2* mutant background (Figure. 4.8). This observation is consistent with a previously described observation, in which the expression pattern of a neomycin resistance gene inserted into the *Hoxa2* locus was shown to be the same as the wild-type *Hoxa2* expression in the hindbrain. In the absence of *Hoxa2* function the neomycin resistance gene expression was not altered (Rijli et al., 1993).

Chapter 5

Regulatory analysis of the Differential expression of the two *Hoxa2* co-paralogous genes in Pufferfish

5.1 Introduction

Most vertebrates, including mouse and chick, have 39 *Hox* genes organized into four clusters (*Hoxa-Hoxd)* on four different chromosomes. These clusters are believed to have evolved by duplication and divergence of an ancestral homeobox gene (Kappen et al., 1989). In invertebrates, *Hox* genes are organized into a single cluster (Krumlauf, 1992). The vertebrate *Hox* genes of each cluster have been aligned by their position within the cluster and their encoded amino acid sequence to the *Hox* genes of *Drosophila*. In this way, *Hox* genes can be organized into 13 paralogous groups.

It has been shown that ray-finned fishes contain extra *Hox* clusters, as compared to other vertebrates (Amores et al., 1998; Amores et al., 2004). For example, zebrafish has a seven *Hox* cluster organization with a total of 51 genes (Amores et al., 1998). Linkage analysis to adjacent non-*Hox* genes and sequence comparison showed that these seven clusters evolved by (whole genome) duplication of the four mammalian *Hox* clusters (Amores et al., 1998). Therefore, each of the four mammalian clusters can be assigned to two duplicates, e.g. the mammalian *Hoxa* cluster is duplicated to two clusters, *Aa* and *Ab*, in fish. One exception is the *HoxD* cluster, where the duplicate (*Hoxdb*) was lost during evolution.

The presence of seven *Hox* clusters has been describe for several other ray-finned fishes, including medaka (*Oryzias latipes*), pufferfish (*Takifugu rubripes*), killifish (*Fundulus heteroclitus*) and striped bass (*Morone saxatilis*) (Aparicio et al., 2002; Misof and Wagner, 1996; Naruse et al., 2000; Snell et al., 1999).

Genome duplication leads to the generation of paralogous groups of the duplicated genes. During evolution, the duplicated gene may lose its function by degeneration to a pseudogene, or it can be lost from the genome; another possibility is that they can gain a new func-

tion (neo-functionalization) or take only part of the function of the original gene; both genes may act complementary to achieve the function of the original gene (sub-functionalization).

One example of sub-functionalization has been shown for the duplicated *Hoxb1a* and *b* genes in zebrafish (McClintock et al., 2001; McClintock et al., 2002). Hoxb1 has been shown to establish the r4 identity in several model systems (Studer et al., 1996). Transgenic analysis in mouse has shown that *Hoxb1* expression is initiated by a conserved retinoic-acid response element (RARE), whose activity is triggered by retinoids (Marshall et al., 1994). This initiated expression becomes restricted to r4, and maintained by an auto-regulatory region which consists of three Hox/Pbx binding sites (Pöpperl et al., 1995).

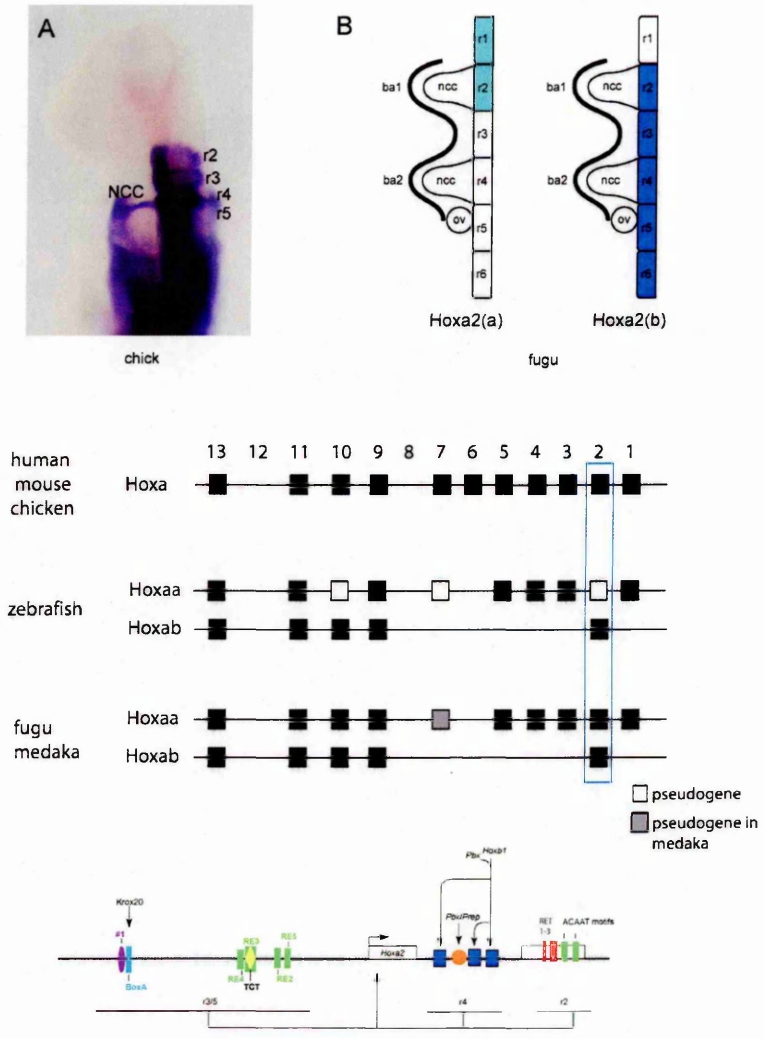


Figure 5.1. Comparison of *Hoxa2* expression and genomic organization of the *Hoxa* clusters between different species.
(A) *In situ* of *Hoxa2* of a chick embryo in the aspect of the hindbrain and BA region. Note the strong expression in rhombomeres 2-5 and in the NCC migrating into the second BA. **(B)** Schematic diagram of the expression domains in the hindbrain of the two fugu co-paralogous genes *Hoxa2(a)* and *(b)*. The *Hoxa2(a)* shows weak expression in r1 and 2, whereas *Hoxa2(b)* shows stronger expression in r2 through r5(Amores et al., 2004).**(C)** Schematic diagram of the *Hoxa* clusters organization in different species including amniotes (human, mouse and chicken), zebrafish, fugu and medaka. In zebrafish, fugu and medaka the *Hoxa* clusters are duplicated. Blue box indicates the position of the *Hoxa2* genes. **(D)** Schematic diagram of the regulatory modules directing *Hoxa2* expression in different rhombomeres.

In zebrafish, *Hoxb1b* is expressed early and transiently in the hindbrain during development, while *Hoxb1a* is expressed strongly in r4 at later stages of development (McClintock

et al., 2001). The expression patterns of *Hoxb1a and Hoxb1b,* together, generate a pattern similar to that of the single mouse *Hoxb1* gene, which is expressed throughout hindbrain development. Interestingly, zebrafish *Hoxb1b* has only the RARE, which initiates early gene expression, whereas the sequences of each of the three Hox/Pbx sites within the autoregulatory region are diverged. In contrast, *Hoxb1a* has lost the RARE, but retained the sequences of the three Hox/Pbx binding sites. This is consistent with the observation that the early expression of *Hoxb1a* is not initiated and it is expressed only at later stages (Prince and Pickett, 2002).

The duplicated *Hoxa2* genes in fugu provide an example of both neo- and sub-functionalization (Figure 5.1). *Hoxa2* is expressed posterior to the r1/2 boundary in the hindbrain of chick, mouse and zebrafish (see Figure 5.1A and also Krumlauf, 1993; Prince and Lumsden, 1994; Wilkinson et al., 1989). The duplicated genes of fugu, *Hoxa2(a)* and *(b)*, show diverged expression from each other (Figure 5.1B and see also Amores et al., 2004). *Hoxa2(a)* shows expression in r1 and r2, whereas *Hoxa2(b)* is expressed from r2 to r5 similar to the expression of *Hoxa2* observed in other modelsystems (Amores et al., 2004). Thus both co-paralogous genes are expressed in r2, while only one is exclusively expressed in r3 through 5 (sub-functionality). The fact that *Hoxa2a* is also expressed in r1, which is not the case in other model systems including zebrafish, suggests that it has gained a new function in r1 (neo-functionality). In medaka, there are also two co-paralogous genes of *Hoxa2*, in contrast to zebrafish, where *Hoxa2(a)* has become a pseudogene (Amores et al., 2004).

Here, I investigate the nature of this differential expression of *Hoxa2(a)* and *(b)* by analyzing the *cis*-regulatory elements controlling rhombomeric expression of both genes. There are three modules which direct *Hoxa2* expression in the hindbrain, including the r3/5 module in the intergenic region, the r4 module located in the intron and the r2 module located in the second exon (see Chapter 3, 4 and Nonchev et al., 1996b).

I cloned the r3/5, r4, and r2 regulatory modules of medaka and fugu, linked them to a *lacZ* reporter gene, and assayed their activity using chick electroporation. I show that subtle sequence drift in the elements of each module of *Hoxa2(a)* have occurred, and these are re-

sponsible for the differential expression of the two co-paralogous genes *Hoxa2(a)* and *(b)*.

5.2 Results

5.2.1 Identification of changes in the *Hoxa2* r3/5 regulatory region leading to differential expression of *Hoxa2(a)* and *(b)*

Fugu *Hoxa2(a)* shows no r3/5 expression in the hindbrain, whereas *Hoxa2(b)* is expressed in these domains (Figure 5.1B and Amores et al., 2004). The *Hoxa2* r3/5 module in mouse consists of several elements embedded in the intergenic region of *Hoxa2/3* (Maconochie et al., 2001; Nonchev et al., 1996b). I cloned the fugu r3/5 enhancers of *Hoxa2(a)* and *(b)* and linked these fragments to a *lacZ* reporter gene. As predicted by *in situ* data, the *Hoxa2(b)* r3/5 enhancer mediates strong expression in the hindbrain, while the *Hoxa2(a)* r3/5 enhancer shows much weaker expression in the hindbrain, mainly in r5 (Figures 5.2B,C). I also confirmed this observation by assaying the medaka r3/5 regulatory regions of *Hoxa2(a)* and

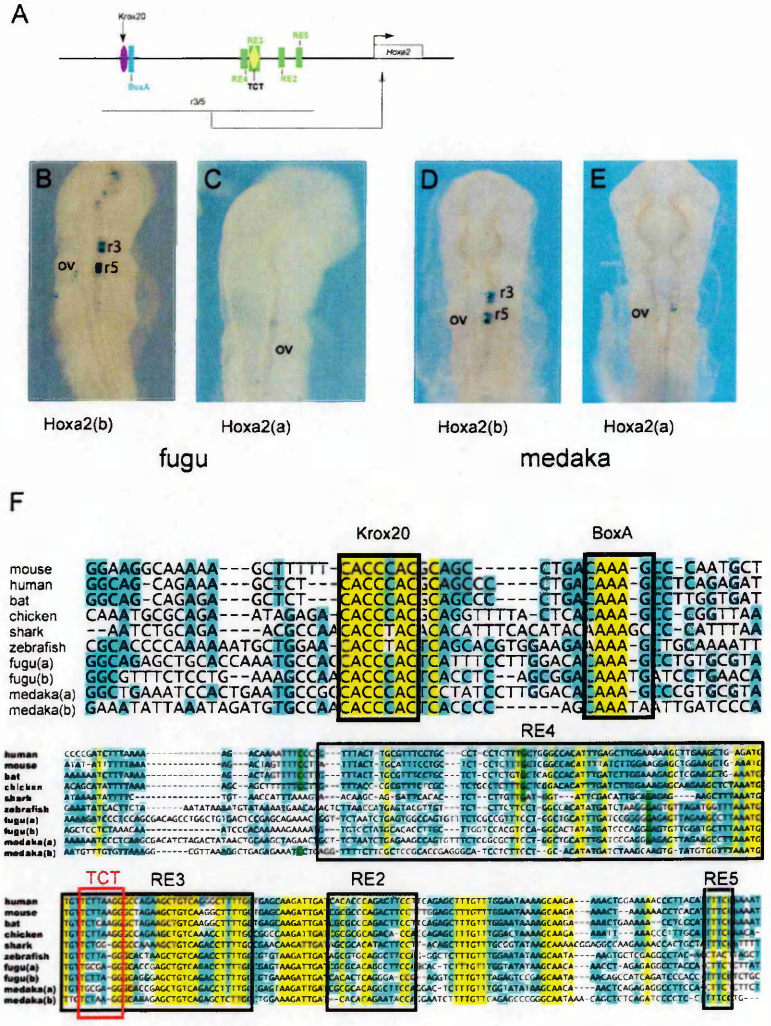


Figure 5.2. Analysis of the fugu and medaka *Hoxa2(a)* and *(b)* r3/5 enhancer activity and sequence alignment of the r3/5 enhancer elements. **(A)** Schematic diagram illustrating the r3/5 module of *Hoxa2*. **(B-E)** Chick embryos electroporated with reporter constructs containing *Hoxa2(a)* and *(b)* r3/5 enhancers from medaka and fugu. Figures **B** and **C** show electroporated embryos with the fugu r3/5 enhancer region of *Hoxa2(b)* and *(a)* respectively. Figures **D** and **E** show electroporated embryos with the medaka r3/5 enhancer region of *Hoxa2(b)* and *(a)* respectively. **(F)** Alignment of the sequence of various species of the r3/5 enhancer. The individual r3/5 elements (Krox20, BoxA, RE4, TCT, RE3, RE2 and RE5) are boxed.

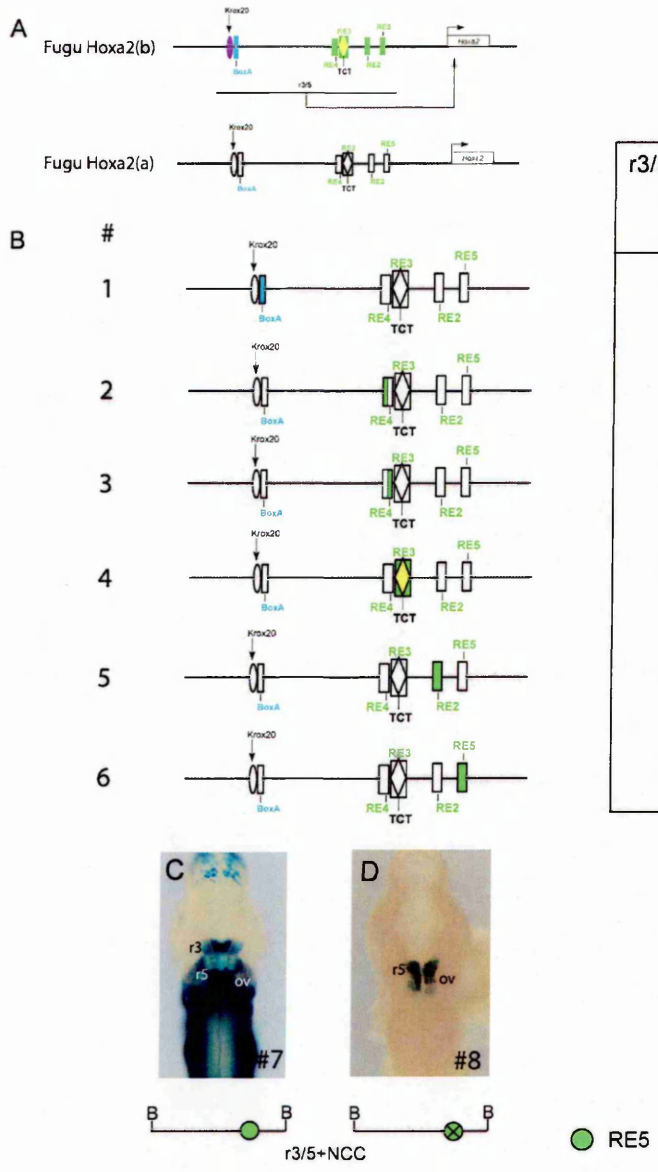
(b) using the chick electroporation assay (Figures 5.2D,E).

Next, I performed an alignment of the previously identified r3/5 elements (Krox20, BoxA, RE4, TCT motif, RE3, RE2) of *Hoxa2* from different species (Figure 5.2F). I observed no diverged sequence in the Krox20 binding region of the *Hoxa2(a)* regulatory region of fugu or medaka. The Krox20 binding motif is highly conserved among all species only shark and zebrafish showing a single nucleotide change at the fifth position. The amniotes, fugu, and medaka have the following Krox20 binding motif: CACCCAC, whereas shark and zebrafish have one nucleotide different: CACCTAC. I therefore focused on the other r3/5 elements and observed slight differences in the medaka and fugu *Hoxa2(a)* and *(b)* regulatory regions. The BoxA motif also shows a high degree of conservation (consensus sequence MAAAKV) between the analyzed species, including the *Hoxa2(a)* and *(b)* r3/5 regions of fugu and medaka. In fact, the sequence of the BoxA motifs of medaka and fugu *Hoxa2(a)* are more similar to the other species than to that of *Hoxa2(b)*. Further, I noted small changes in the spacing between the Krox20 binding site and the BoxA motifs between medaka and fugu *Hoxa2(a)* and *(b)*.

The RE4 element shows extensive sequence variation between species, which makes it difficult to pinpoint the important changes between *Hoxa2(a)* and *(b)*. The most notable changes were in the TCT motif, which is embedded in the RE3 element. The first three highly conserved nucleotides, TCT, have evolved to TGC in the r3/5 regulatory module of both fugu and medaka *Hoxa2(a)*. The RE2 motif does not exhibit any notable changes in fugu and medaka between *Hoxa2(a)* and *(b)*.

Further, I observed a diverged sequence in a conserved region which has been not described before. This region is located 3' of the RE2 element and has the consensus sequence TTTCC in mammals, chick, fugu(b), and medaka(b). Zebrafish was the only species in which this sequence was not present. This element shows diverged changes in fugu(a) and medaka(a) to **C**TTC**T**. I termed this element rhombomeric element 5 (RE5).

In order to experimentally test which of the observed changes are able to contribute to the differential expression of *Hoxa2(a)* and *(b)*, I designed various constructs where either



r3/5 expression efficiency	
n	%
16	0
25	32
18	0
18	67
31	6
33	33

Figure 5.3. Identifying the r3/5 elements responsible for differential expression of fugu *Hoxa2(a)* and *(b)*.
(A) Schematic diagrams of the r3/5 modules of fugu *Hoxa2(b)* and *(a)*. Note that the *Hoxa2(b)* elements are colored, whereas the elements of *Hoxa2(a)* are white. **(B)** Schematic diagrams of the construct used for electroporation and regulatory analysis of each element of the *Hoxa2(a)* and (b) r3/5 elements. Colored elements illustrate that this element has been changed from *Hoxa2(a)* to *Hoxa2(b)*. Each modified element has been linked to a *lacZ* reporter gene. The numbers on the right of each construct shows the numbers of embryos electroporated and how many of them showed r3/5 restricted expression. The numbers in brackets represent the numbers of embryos showing r3/5 expression in percentage. **(C)** Deletion analysis of the RE5 element in transgenic mice. The left 9.5 dpc embryo is transgenic with the WT *Hoxa2*

parts or the whole sequence of one element of the r3/5 enhancer of fugu *Hoxa2(a)* is changed to the sequence of *Hoxa2(b)* (Figure 5.3B). In the first construct, I changed the sequence of the BoxA element, as well as the spacing between BoxA and the Krox20 motif (construct #1). This construct, and the following constructs, were linked to *lacZ* reporter genes and assayed using chick electroporation. None of the embryos electroporated with this construct showed r3/5 mediated reporter expression. This suggests that the subtle changes in the BoxA motif are not primarily responsible for the differential expression patterns of *Hoxa2(a)* and *(b)* of medaka and fugu.

Next, I altered the *Hoxa2(a)* sequence to *(b)* in the RE4 motif sequence. I built two constructs for this purpose; in the first one (construct #2), I changed the first half of the RE4 motif, and in the second construct (#3), I altered the second half of the RE4. Simply changing

the first half of the RE4 element restored function of the fugu *Hoxa2(a)* r3/5 enhancer to some degree; 32% of the electroporated embryos showed specific reporter staining in r3/5. Changes in the second half of the RE4 motif had no effect on the r3/5 enhancer activity.

In another construct (construct #4), I changed the RE3 element, which includes the TCT motif. The majority of embryos (67%) electroporated with this construct exhibit r3/5 specific *lacZ* expression in the hindbrain. These changes in the sequence of the RE3 play a major role in the different r3/5 enhancer activity of *Hoxa2(a)* and *(b)* in medaka and fugu. Next, I tested the R2 motif, but alterations in the sequence of this element from *Hoxa2(a)* to *(b)* do not exhibit an increase of r3/5 enhancer activity (construct #5).

Finally, I tested whether changes in the newly identified r5 element, RE5, are able to restore r3/5 activity (construct #6). 33% of the embryos electroporated with this construct showed r3/5 specific reporter staining, suggesting that this element not only has an important role in the r3/5 module, but also that its diverged sequence accounts for the differential expression of *Hoxa2(a)* and *(b)* in fugu and medaka. In order to confirm that this element has an important role in the r3/5 module, I deleted this element from the mouse *Hoxa2 BglII* fragment, linked it to a *lacZ* reporter gene, and tested it in transgenic mouse embryos (Figures 5.3C, D). While the control fragment directed r3/5 and NCC reporter staining (construct #7), the mutated fragment showed only expression in r5 and weak expression in r6 (construct #8).

5.2.2 Identification of changes in the *Hoxa2* r4 regulatory region leading to differential expression of *Hoxa2(a)* and *(b)*

Hoxa2(a) is not expressed in r4, whereas *Hoxa2(b)* is strongly expressed in r4 (Figure 5.1B and Amores et al., 2004). I have shown that the r4 enhancer of *Hoxa2* contains four components embedded within the *Hoxa2* intron (see Chapter 3 and Figure 5.4A). In order to explain the differential expression of *Hoxa2(a)* and *(b)* in r4, I cloned the *Hoxa2(a) and (b)* intronic regions, containing the r4 regulatory elements from both medaka and fugu. I linked each fragment to a *lacZ* reporter gene and assayed for r4 enhancer activity using chick electro-

poration. Consistent with *in situ* data, the *Hoxa2(b)* r4 regulatory regions direct stronger reporter expression in r4 than does the *Hoxa2(a)* r4 regulatory region, although the medaka

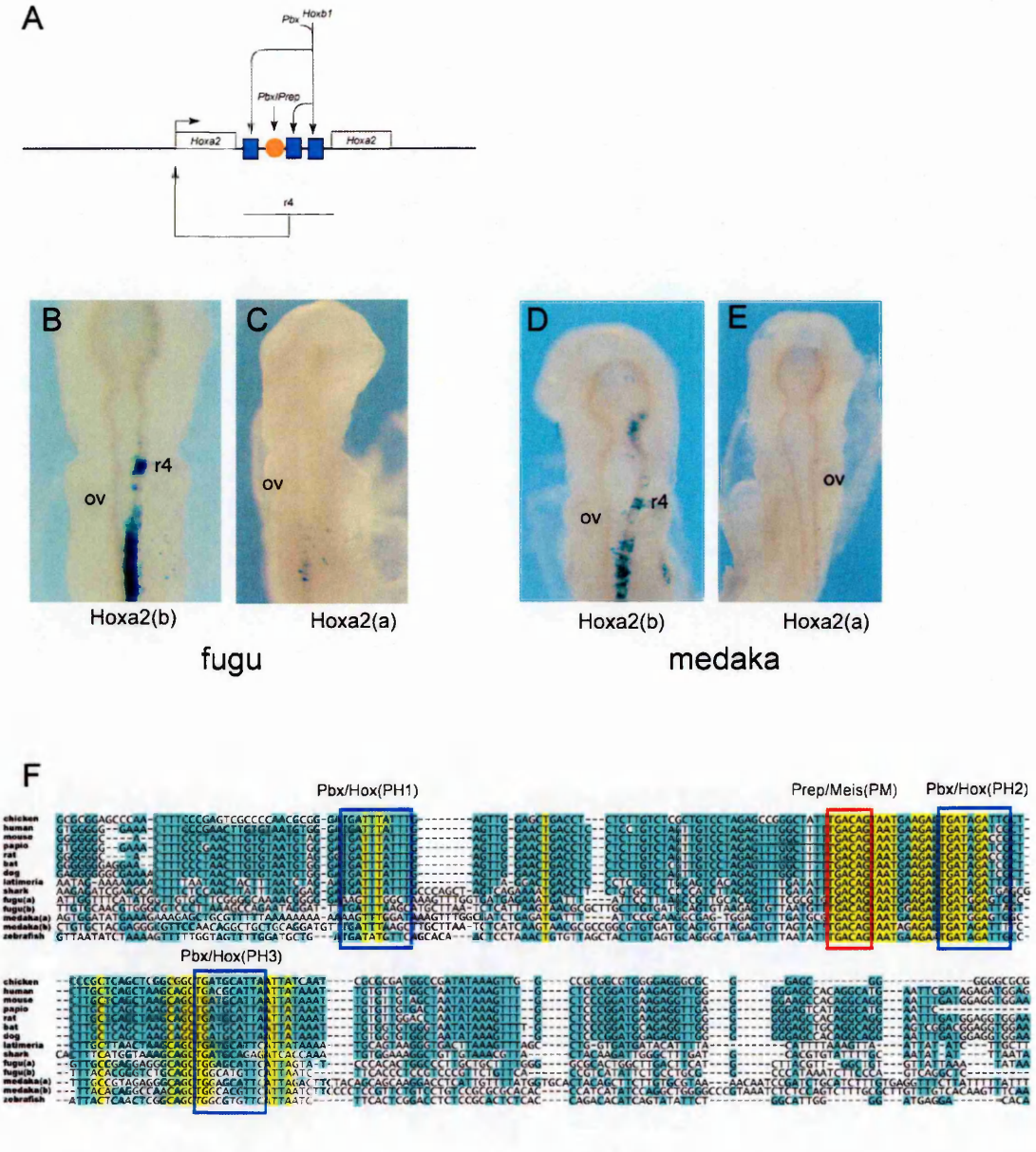


Figure 5.4. Analysis of the fugu and medaka *Hoxa2(a)* and *(b)* r4 enhancer activity and sequence alignment of the r4 enhancer elements.
(A) Schematic diagram of the *Hoxa2* r4 module. **(B, C)** Electroporated embryos with reporter construct linked with fugu *Hoxa2* r4 enhancer from *Hoxa2(b)* and (a) respectively. Figures **D** and **E** show chick embryos electroporated with reporter construct fused with the medaka r4 enhancers from *Hoxa2(b)* and *(a)* respectively. **(F)** Sequence alignment of the *Hoxa2* r4 module with its element (PH1, PM, PH2, PH3) boxed. r4, rhombomere 4; ov, otic vesicle.

Hoxa2(b) module was remarkably weaker than the fugu *Hoxa2(b)* module (Figures 5.4B-E).

Next, I performed an alignment of the r4 module which includes three Hox/Pbx (PH-1-3) elements and a Prep/Meis (PM) element (Figure 5.4F). The first and third elements show a high degree of diverged sequence between *Hoxa2(a)* and *(b)*. In particular, the second G in the consensus sequence, TGATNNATGC, is important for exd/pbx binding (Chan et al., 1997;

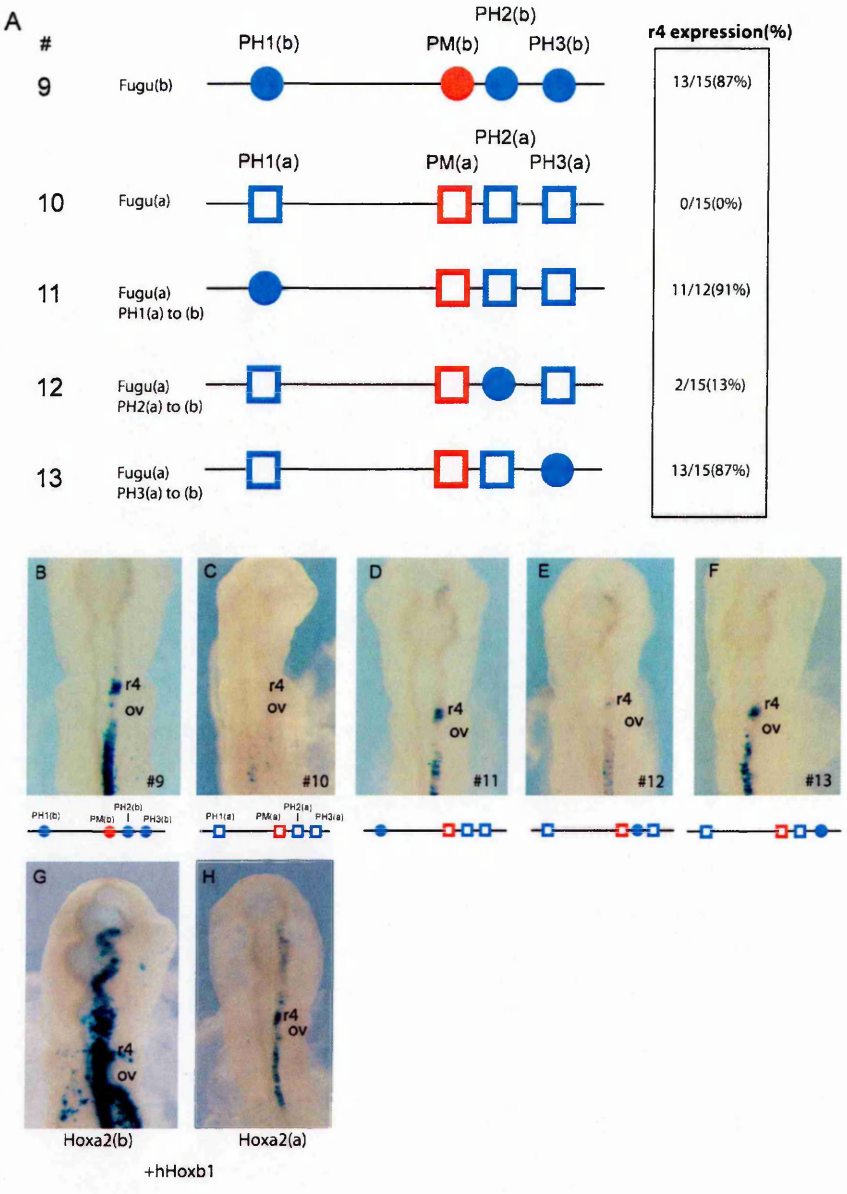


Figure 5.5. Identifying the r4 elements responsible for differential expression of fugu *Hoxa2(a)* and *(b)*.. **(A)** Schematic diagrams of the constructs used for identifying the important changes responsible for differential r4 enhancer activity of fugu *Hoxa2(a)* and *(b)*. In circles the r4 elements of fugu *Hoxa2(b)* are represented, whereas the ones of fugu *Hoxa2(a)* are illustrated in squares. **(B-F)** Represented chick embryos electroporated with the above listed constructs. **(G, H)** Hoxb1 *trans*-activation of the fugu *Hoxa2(b)* and *(a)* r4 enhancers. r4, rhombomere 4; ov, otic vesicle.

Chan and Mann, 1996). This G has evolved to an A in the first bipartite Hox/Pbx (PH1) binding site of *Hoxa2(a)* in fugu and medaka, giving the consensus sequence **A**AGTTTGGCMA (Figure 5.4).

I then asked whether I could restore r4 expression of *Hoxa2(a)* by converting each of the elements to the *Hoxa2(b)* sequence (Figure 5.5A). Restoration of r4 activity was most significant when the first or third PH(a) sites to PH(b) were changed (constructs #11 and #13). In

contrast, when the second site of *Hoxa2(a)* was changed to (b) (construct #12), r4 activity was not restored (construct #10).

I have shown that ectopic *Hoxb1* expression can *trans*-activate the chick *Hoxa2* r4 enhancer (see Chapter 3). I, therefore, asked whether overexpressed Hoxb1 is capable of *trans*-activating the fugu *Hoxa2(a)* enhancer, and whether it is different from the fugu *Hoxa2(b)* *trans*-activation (Figures 5.5G, H). As shown in figure 5H, overexpression of human Hoxb1 under the control of the CMV promoter, is capable of *trans*-activating the fugu *Hoxa2(a)* enhancer, showing strong expression in r4, and also in the adjacent rhombomeres. By this way I was able to restore the r4 *Hoxa2(a)* enhancer activity. I also *trans*-activated the fugu *Hoxa2(b)* with Hoxb1 and observed a much stronger *trans*-activation in the hindbrain, compared to the *trans*-activation of the fugu *Hoxa2(a)*.

5.2.3 Identification of the changes that have occurred in the *Hoxa2* r2 regulatory region leading to differential expression of *Hoxa2(a)* and *(b)*

In situ analysis of fugu *Hoxa2(a)* and *(b)* shows that both genes are expressed in r2, although the expression of *Hoxa2(a)* is weaker in r2 than is the expression of *Hoxa2(b)(Amores et al., 2004)*. The r2 module of *Hoxa2* consists of five elements (RTE1-3 and ACAAT 1-2) embedded in the second exon (Figure 5.6A). I cloned the r2 modules of *Hoxa(a)* and *(b)* from fugu and medaka and linked them to the *lacZ* reporter gene and evaluated their enhancer activity in the chick electroporation assay. As predicted by *in situ* data, the r2 enhancers of fugu and medaka *Hoxa2(b)* mediate stronger reporter expression than the r2 enhancers of *Hoxa2(a)* which direct consistently weaker expression (Figures 5.6B-E).

Sequence alignments of the *Hoxa2* r2 modules revealed that the ACAAT motifs, which I have shown to initiate r2 expression, are highly conserved and do not show any sequence divergence between *Hoxa2(a)* and *(b)* (Figure 5.6F).

The RTE motifs do show diverged sequences in the elements of RTE1 and RTE2. To investigate whether these changes account for the differential expression, I built an r2 enhancer construct (construct #16) which contains a region of fugu *Hoxa2(a)* at its 5' end and a region of fugu *Hoxa2(b)* at its 3' end (Figure 5.7). The other construct (#17) was built in the reciprocal way. It has the fugu *Hoxa2(b)* fragment at its 5' end and a region of fugu *Hoxa2(a)* at its 3' end (construct #16). The two portions are linked by a conveniently conserved *HincII* site, which is present in both *Hoxa2(a)* and *(b)* r2 modules. Therefore, I did not change any sequence at the linking point. The first construct gave strong r2 expression, whereas the r2 enhancer activity of #17 construct was much weaker. This supports our idea that the changes responsible for the different r2 enhancer activity between *Hoxa2(a)* and *(b)* lie in the region which contains the elements (RTE1-3) that maintain r2 expression.

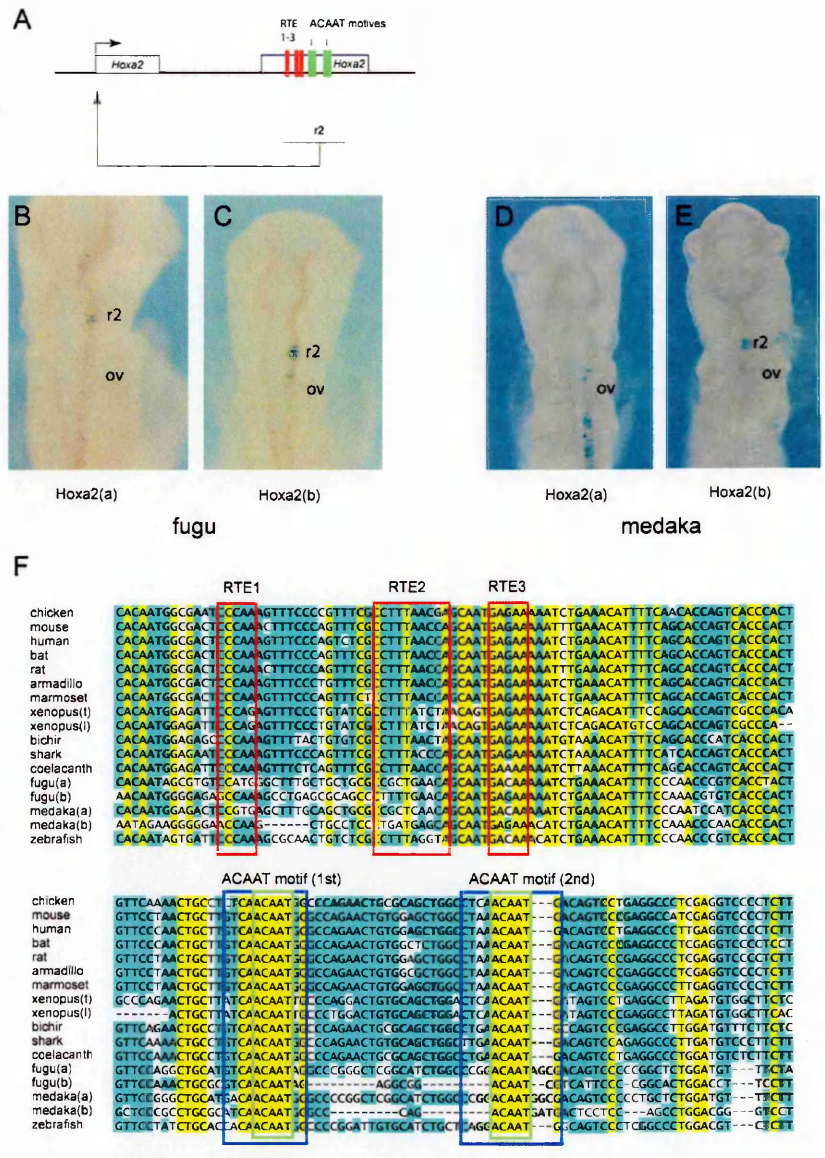


Figure 5.6. Analysis of the fugu and medaka *Hoxa2(a)* and *(b)* r2 enhancer activity and sequence alignment of the r2 enhancer elements. **(A)** Schematic diagram of the *Hoxa2* r2 module. **(B-E)** Electroporated embryos stained for reporter expression. The embryos were electroporated with a *lacZ* reporter constructs fused with the fugu **(B,C)** and medaka **(D, E)** r2 enhancer region from *Hoxa2(a)* and *(b)*. **(F)** Sequence alignment of the r2 module with RTE (1 to 3) elements and the two ACAAT motifs boxed. r2, rhombomere 2; ov, otic vesicle.

In order to identify the important changes, I used construct #15 and changed the sequence from *Hoxa2(a)* to *(b)* throughout the first portion of the *HincII* fragment (constructs

18-25). Change #20 gave an increase in r2 reporter staining (from 7% to 51%), whereas the other constructs did not exhibit any increase in r2 mediated reporter staining. I then changed regions C and D in the same construct (#23) and restored activity from 7% to 70%, which approximates the r2 activity of the r2 fugu *Hoxa2(b)* (construct #15). Since these changes are in the region where the RTE1 and 2 elements are located, I changed the specific sequence of the RTE1 and 2 elements from *Hoxa2(b)* to *(a),* as shown in construct #24. I was able to convert the strong r2 enhancer activity of 78% to 49%. In the reciprocal experiment, I changed the sequence of the RTE1 and 2 elements from *Hoxa2(b)* to *(a),* and I was able to restore the r2 activity of fugu *Hoxa2(a)* to almost 50%. The sequence of RTE1 has a CCA core in all analyzed species, including *Hoxa2(b)* of medaka and fugu, whereas, in medaka and fugu *Hoxa2(a)*, the sequence of this core is changed to CCA (fugu) or CGT (medaka). Together these results indicate that the changes in RTE1 play an important role in the differential activity of the fugu *Hoxa2(a)* and *(b)* r2 enhancers.

5.3 Discussion

In this study, I investigated the basis of the differential expression of the duplicated genes, *Hoxa2(a)* and *(b),* of fugu and medaka (Figure 5.8). I used sequence comparison and the chick electroporation assay to analyze the differential activity of the rhombomeric modules (r2, r3/5 and r4) of the two co-paralogous genes *Hoxa2(a) and (b).*

I performed sequence alignments of known regulatory regions directing rhombomeric expression of *Hoxa2* and identified in each of the medaka and fugu regulatory modules of *Hoxa2(a)* sequence drift. I assayed whether these nucleotide changes were responsible for differential expression, by designing constructs in which the sequence of each of these elements was changed from *Hoxa2(a)* to *(b)* and *vice versa*. Next, I tested them in the chick electroporation system and observed, in several cases, that the enhancer activity was changed. These changes were located in elements which have been shown to have an important role in regulat-

ing *Hoxa2* expression in the hindbrain.

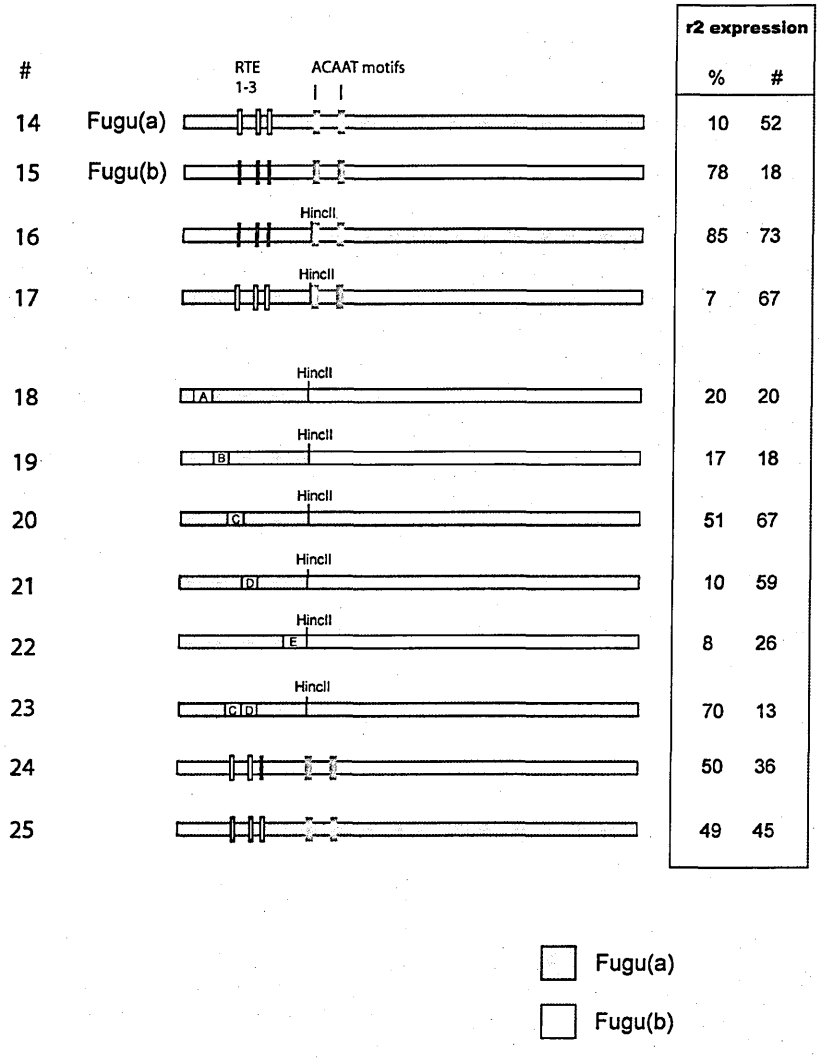


Figure 5.7 Identifying the r2 elements responsible for differential expression of fugu *Hoxa2(a)* and *(b)*.
Constructs used for identifying the regions responsible for differential expression in r2 of fugu *Hoxa(a)* and *(b)*. The grey region represents region where the sequence has been changed from *Hoxa2(b)* to *Hoxa2(a)* and the white region (including segments A-E in constructs 18-23) illustrates the sequence from the *Hoxa2(b)* region.

I identified in each of the modules several changes. In the r3/5 enhancer I observed sequence drift in the RE3 element, which contains a TCT motif known to be important for r3/5 enhancer activity (Maconochie et al., 2001). Changing the sequence of this element from *Hoxa2(a)* to *(b)* had a profound effect on the enhancer activity. However, other elements known to be important for r3/5 activity, like the Krox20 binding site and the BoxA element of the r3/5 module, showed no sequence drift.

In the *Hoxa2(a)* r4 enhancers of fugu and medaka, I identified sequence changes in the first and third Hox/Pbx sites. I showed that these changes are contributing to the differential expression of *Hoxa2(a)* and *(b)* in fugu. Interestingly, overexpression of Hoxb1 in the hindbrain was able to restore r4 enhancer activity and, in addition, led to *trans*-activation of the fugu *Hoxa2(a)* enhancer in adjacent rhombomeres. Hoxb1 may bind to the r4 elements of the fugu *Hoxa2(a)* enhancer, but the affinity maybe weaker, and result in the lower level of *trans*-activation, as compared to that of *Hoxa2(b)*.

Both genes, *Hoxa2(a)* and *(b)*, show expression in r2, although the expression of *Hoxa2(b)* is much stronger than that of *(a)*. The r2 enhancer consists of several elements including two ACAAT motifs and three RTE motifs. Comparing the sequence of these elements from the two fugu and medaka *Hoxa2* co-paralogous genes with the sequence of other species, I observed no changes in the highly conserved ACAAT motifs between *Hoxa2(a)* and *(b)*. Although the spacing between the two ACAAT motifs was much reduced in *Hoxa2(b)* of both medaka and fugu, as compared to that of *Hoxa2(a)*. However, the region between the ACAAT motifs contains no regulatory elements and I have shown that the spacing between the two ACAAT motifs is not important (see Chapter 5). The only significant changes I observed were in the RTE1 and 2 elements, and, in fact, changes in the region of the RTE1 of the *Hoxa2(a)* enhancer to *Hoxa2(b)* restored r2 activity. Based on these findings, I am able to explain the basis of the differential expression of *Hoxa2(a)* and *(b)*.

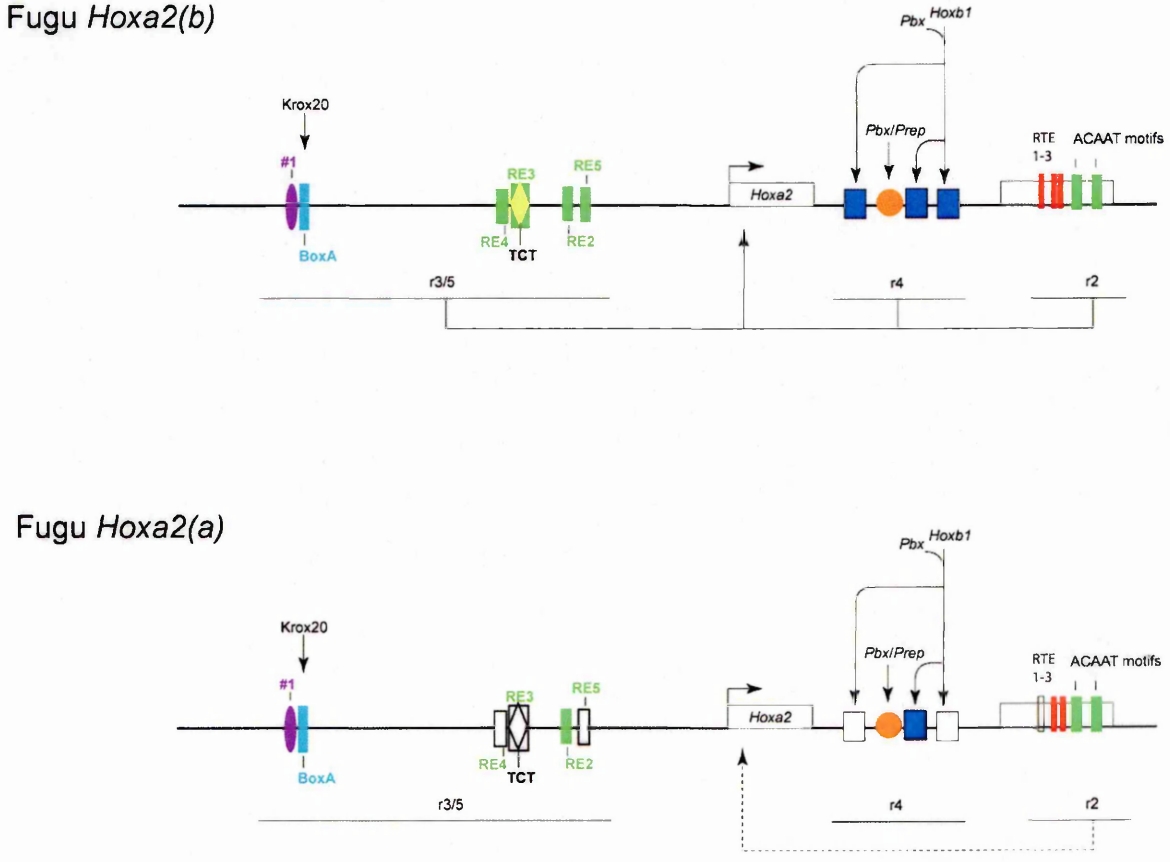


Figure 5.8. Summary of the rhombomeric regulatory modules of fugu *Hoxa2(a)* and *(b)*.
The top schematic diagram illustrates the fugu *Hoxa2(b)* module and the fugu *Hoxa2(a)* module is illustrated on the bottom. The white squares in the fugu *Hoxa2(a)* module illustrate the inactive element whereas the colored ones are active. The dotted line in the fugu *Hoxa2(a)* module illustrates that the r2 is partially active (see Figure 5.1B).

However, fugu *Hoxa2(a)* shows additional expression in r1, which is not the case for *Hoxa2(b)* or other analyzed model systems, including mouse, chick or zebrafish. In our chick electroporation experiments, I did not observe any r1 reporter expression of any of the tested *Hoxa2(a)* constructs. It is therefore possible that a new regulatory module directing *Hoxa2* in r1 evolved in other regions. It has been shown that the fibroblast growth factor-8 (FGF8) normally represses *Hoxa2* expression in r1 (Irving and Mason, 2000). Thus, it is also possible that this repression mechanism has changed in *Hoxa2(a)*, either in a direct or in an indirect manner.

In most cases, the changes in the regulatory elements, which led to differential expression of *Hoxa2(a)* and *(b),* were the same in both medaka and fugu. This implies that the sequence drift in the regulatory region of *Hoxa2(a)* occurred before the evolutionary split leading to the lineages of the spiny ray fishes, medaka and fugu. Interestingly, in the zebrafish lineage the *Hoxa2(a)* became a pseudogene.

Chapter 6

Krox20 represses *Hoxb1* expression by a direct and highly conserved mechanism in rhombomeres 3 and 5

6.1.1 Introduction

Hoxb1 expression is initiated by retinoids, showing at early stages a broader domain of expression in the hindbrain, but its expression becomes restricted to r4 during the course of segmentation. The initiated expression is maintained by an autoregulatory loop, and also through Hoxa1 and Hoxb2. Previous studies have shown that several molecules are involved in restricting *Hoxb1* expression. For example, a conserved retinoic response element in the *Hoxb1/b2* intergenic region has been shown to mediate downregulation of the expression of *Hoxb1* in r3 and r5 (Studer et al., 1994). Removal of this element leads to expansion *Hoxb1* expression into adjacent rhombomeres (Studer et al., 1994). In zebrafish, the gene *variant hepatocyte nuclear factor 1* (*vhnf1*) has been shown to downregulate of *Hoxb1* expression in r5 and r6 (Gaufo et al., 2003; Wiellette and Sive, 2003), and other studies have shown that *Hox3* genes repress *Hoxb1* expression in r6 (Gaufo et al., 2003).

Krox20 has been implicated in downregulation of *Hoxb1* expression in the hindbrain of chick embryos (Giudicelli et al., 2003; Mechta-Grigoriou et al., 2000). Two members of the NGFI-A binding corepressors, Nab1 and Nab2, have been shown to interact with Krox20, and repress the transcription of its target genes (Russo et al., 1995; Svaren et al., 1996; Swirnoff et al., 1998). *Nab1* and *Nab2* are both expressed in r3 and r5 and it has been shown that their expression is Krox20 dependent (Mechta-Grigoriou et al., 2000). This implies that Krox20 can not only activate expression in r3 and r5, as shown e.g. for *Hoxa2*, *Hoxb2* and *EphA4* (Nonchev et al., 1996b; Sham et al., 1993; Theil et al., 1998) but, it is also able to act as a repressor in these domains, via interactions with its co-repressors.

In this study, I used sequence comparison among different species in the intergenic region of the *Hoxb1/b2* intergenic region in order to identify putative Krox20 binding sites. I identified one highly conserved region in a regulatory region, which has been shown to exhibit r3/5 repressor activity. This highly conserved region is very similar to the Krox20 consensus binding site. I show that this highly conserved region can bind Krox20 *in vitro*, and that this region does mediate repressor activity in the hindbrain. Deletion of this binding site in a transgenic construct leads to an expansion of reporter expression in r3/5. This work has been performed in collaboration with Dr. Michele Studer, Charles Banks and Dr. John McCarthy.

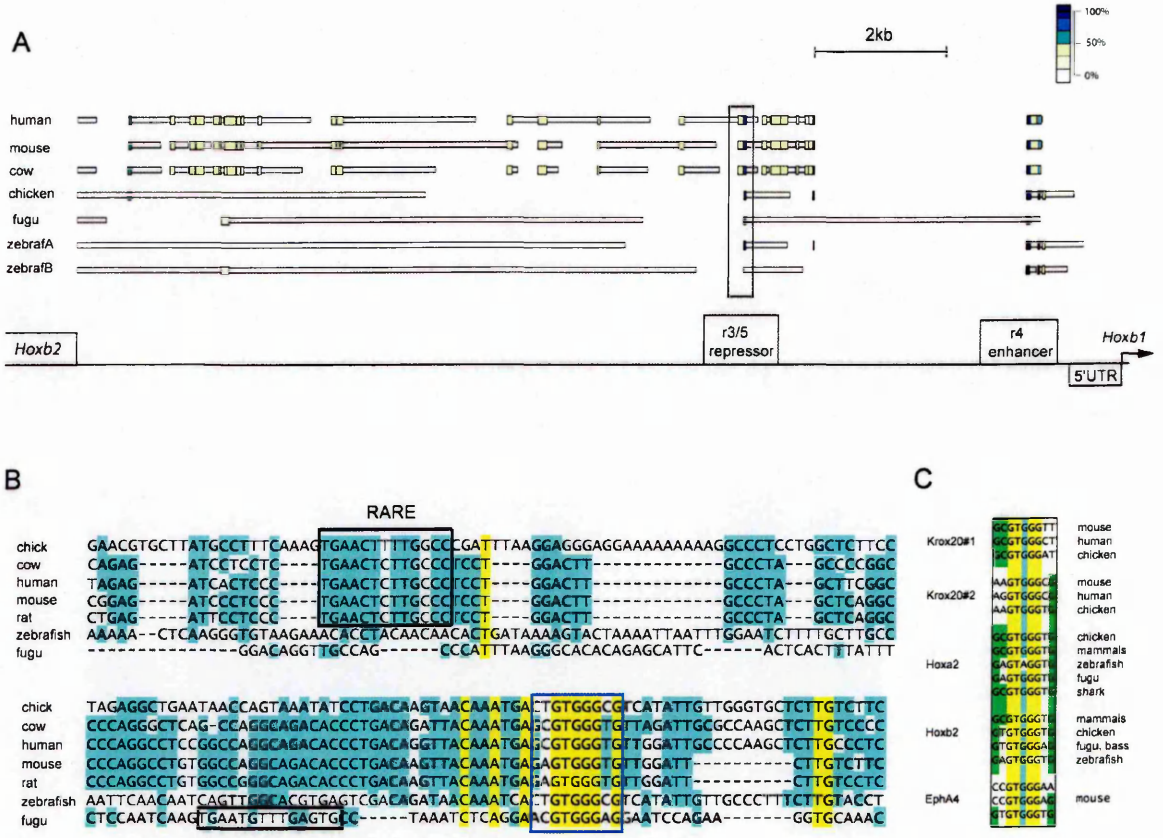


Figure 6.1. Global Sequence Alignment of the *Hoxb1/2* intergenic region and local alignment of the *Hoxb1* r3/5 repressor region.
(A) MACAW alignment of the intergenic region of *Hoxb1/2* of different species, including human, mouse, cow, chicken, fugu, zebrafish(a) and zebrafish(b). The color legend on the top right illustrates the color coding in relation to the degree of percent conservation in the MACAW alignment. The legend on the bottom of the alignment shows the relative positions of the *Hoxb1* r4 enhancer, r3/5 repressor region, and the 3' region of the *Hoxb2* gene. **(B)** Local alignment of the *Hoxb1* r3/5 repressor region of different species, including chick, cow, human, mouse, rat, zebrafish, and fugu. The black boxes encompass the RARE sequences, whereas the blue box contains the putative Krox20 binding site. **(C)** Comparison of known Krox20 binding sites from different regulatory regions, including the *Krox20*, *Hoxa2*, *Hoxb2*, and *EphA4* enhancers and of different species, including various mammals, chicken, zebrafish, fugu, shark, and bass.

6.2 Results

6.2.1 Alignment of the intergenic region of *Hoxb1/2* revealed a putative Krox20 binding site

In order to examine whether *Hoxb1* is directly downregulated by Krox20, I compared the intergenic region of *Hoxb1/b2* of different species, including human, mouse, cow, chicken, fugu, zebrafish(a) and zebrafish(b) (Figure 6.1A). I identified three conserved regions within this area (Figure 6.1A). The first region is located immediately upstream of the *Hoxb1* gene, encompasses the auto-regulatory region of *Hoxb1,* and is conserved among all analyzed species. The second region contains the r3/5 repressor element. Here the alignment shows conservation among all species, except zebrafish(b). The third conserved region is located more 5' and shows sequence conservation among mammals, fugu, and zebrafish(b).

I then performed local alignments of these regions and searched for putative Krox20 binding sites. Only the second highly conserved region contained a conserved region with a putative Krox20 binding site (Figure 6.1B). This finding is very intriguing, since an RARE element in this region has been shown to have r3/5 repressor activity. The putative Krox20 site has a highly conserved core sequence, GTGGG, and adjacent sequences have more variation (Figure 6.1B).

In order to analyze these sequence variations, I compared known Krox20 binding sites from different enhancers, including *Krox20, Hoxa2, Hoxb2* and *EphA4,* and from different species including different mammals, chicken, zebrafish, fugu, and bass (Figure 6.1C). The first nucleotide of the putative *Hoxb1* Krox20 site is either a G, C, or A. This is exactly the same variation in the known Krox20 binding sites. The second nucleotide in the *Hoxb1* Krox20 binding site is either a C, T or A. At this position the known Krox20 sites also show a high degree of variation, accepting all nucleotides. The next nucleotides encompass the highly conserved GTGGG core. Here the known Krox20 binding sites also show the same core with a high degree of conservation, except for the Krox20 binding site of zebrafish in the r3/5 *Hoxa2*

enhancer, which has an A instead of a G at the fifth poison. At the 8th position the *Hoxb1* Krox20 bind site has a T, C, or an A at this position. This is the same variation of nucleotides in the known Krox20 binding elements. The last position of the *Hoxb1* Krox20 binding site is conserved with a G. This is also the predominant nucleotide in the known Krox20 binding sites. This comparison between the sequences of the putative *Hoxb1* Krox20 binding site and the sequences of known Krox20 binding sites shows that the variations are remarkably similar.

6.2.2 Krox20 protein interacts with the putative Hoxb1 Krox20 binding site *in vitro*

In order to test for direct interaction between Krox20 protein and the *Hoxb1* Krox20 binding site, I performed electrophoretic mobility shift assay (EMSA) experiments (Figure 6.2). I used labeled double-stranded oligonucleotides spanning the putative *Hoxb1* Krox20

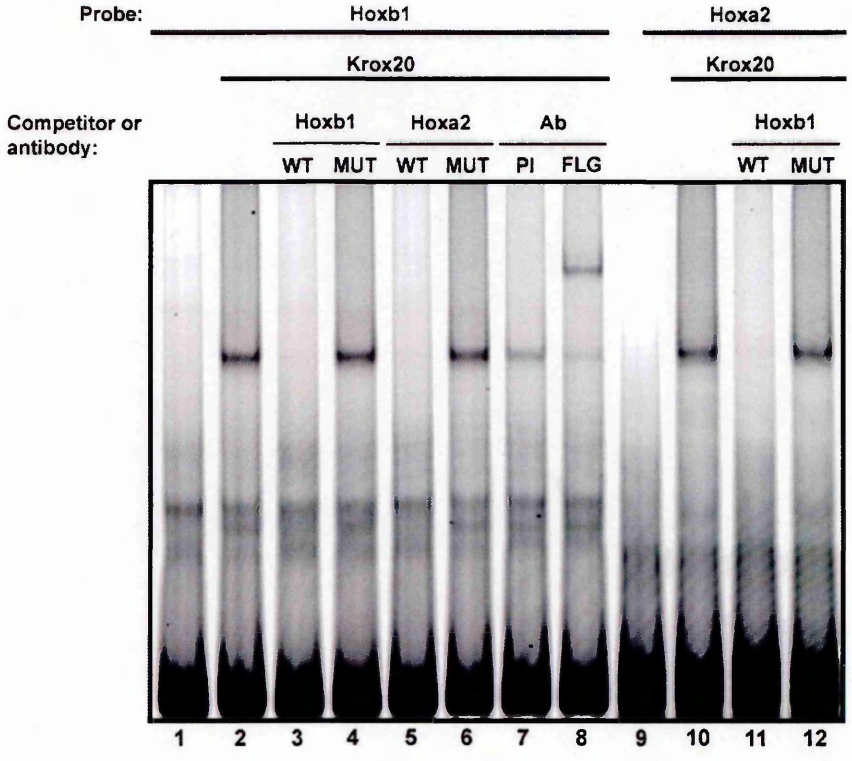


Figure 6.2. Electrophoretic mobility shift assay of binding formation between Krox20 Protein and Hoxb1 sequences.
(Lanes 1 and 9) Control lanes containing either the *Hoxb1* (lane 1) and *Hoxa2* (lane 9) double-stranded oligonucleotides in the absence of Krox20 protein.
(Lanes 2-8) A double-stranded oligonucleotide containing the putative *Hoxb1* Krox20 binding site was incubated with recombinant Krox20 protein in the presence or absence of various competitors and Flag antibody.
(Lanes 10-12) Control gel shift assay where oligonucleotides containing the known *Hoxa2* Krox20 binding site have been mixed with Krox20 protein in the presence or absence of different competitors. EMSA performed by Dr. McCarthy.

binding site as a substrate for complex formation in the presence or absence of different competitors and antibodies (lane 1- 8). Lane 2 shows a slowly migrating band, which disappeared when an excess of different WT competitors were added to the reaction mixtures (lanes 3 and 5). I used both oligonucleodites spanning the putative *Hoxb1* Krox20 bind site (lane 3) and the known Krox20 binding site located in the *Hoxa2* enhancer (lane 5). The band remains when an excess of different mutated competitor oligonucleodites were used (lanes 4 and 6). Again, I used oligonucleotides containing mutated *Hoxb1* Krox20 binding sites (lane 4) and mutated Krox20 binding site of the *Hoxa2* enhancer (lane 6). Since the recombinant Krox20 protein is Flag tagged, I was able to supershift the band by adding anti-Flag antibodies. To further confirm that the slowly migrating band corresponds to Krox20 complex formation, I also used oligonucleotides containing the known *Hoxa2* Krox20 binding site as a control (lanes 9-12). The level of migration of these bands is the same as that obtained with the *Hoxb1* Krox20 binding sites. This strongly suggests that Krox20 protein is able to interact with the *Hoxb1* Krox20 binding site and that this binding is site-specific.

6.2.3 The Krox20 binding site has repressor activity *in vivo*

Next I wanted to evaluate the *in vivo* role of the *Hoxb1* Krox20 binding site, especially whether it exhibits repressor activity. The repressor activity of the *Hoxb1* RARE has been tested by linking various forms of the repressor region to the *Hoxb2* r3/5 enhancer (Studer et al., 1994). These constructs were linked to *lacZ* reporter genes and assayed in transgenic mouse embryos (Studer et al., 1994). I used the same assay system for the Krox20 binding site (Figure 6.3). Construct #5 contains the *Hoxb2* r3/5 enhancer linked to the *lacZ* reporter gene as a control. All of the transgenic animals carrying this construct show r3/5 restricted reporter expression (n=6).

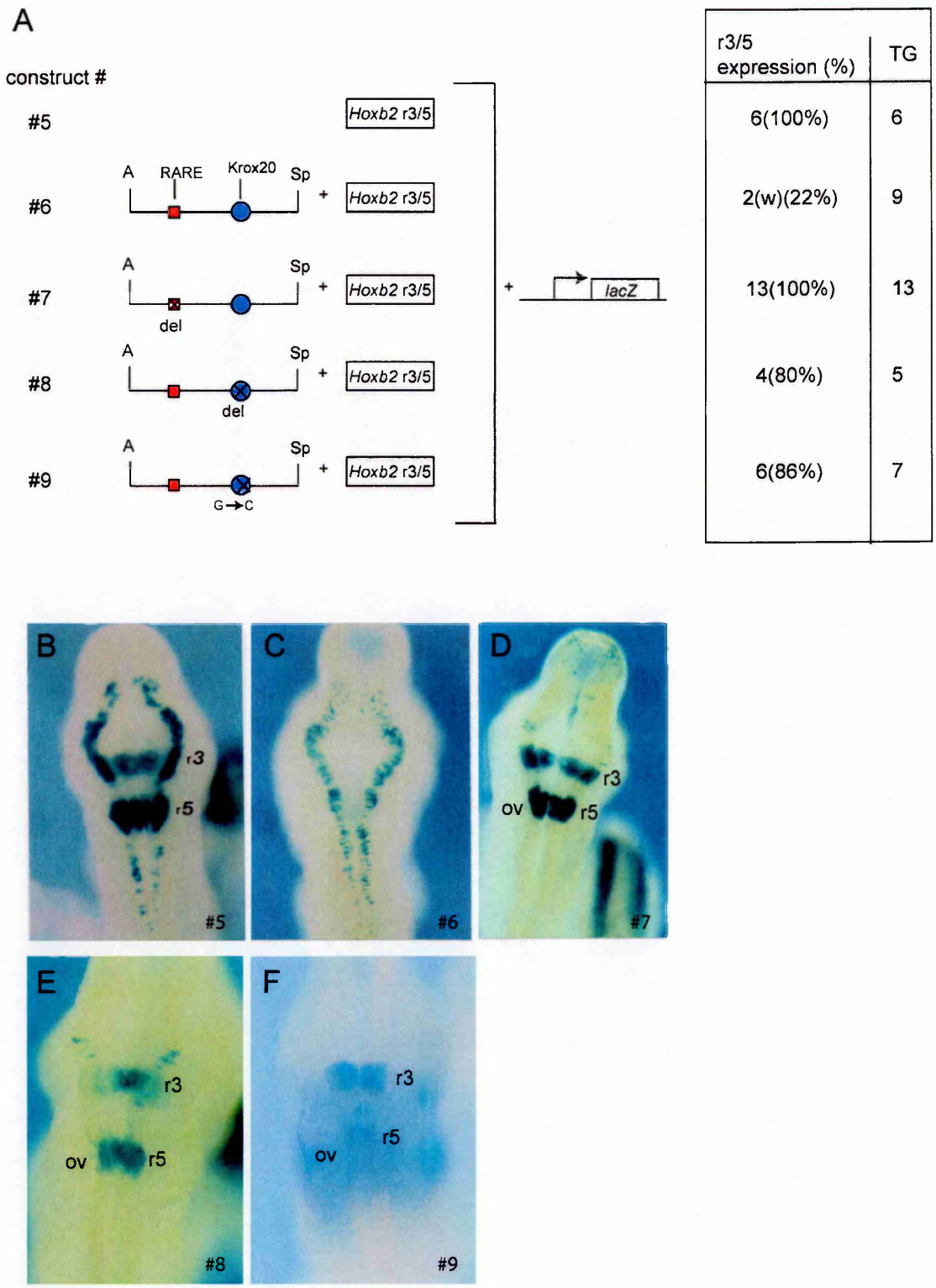


Figure 6.3. Constructs used to assay the repressor activity of different elements of the *Hoxb1* r3/5 region. **(A)** Construct #5 contains only the r3/5 enhancer of *Hoxb2*, whereas in constructs #6 to #9, the *Hoxb2* r3/5 enhancer has been linked to various *Hoxb1* r3/5 repressor regions. All constructs are fused with *lacZ* reporter genes. On the left, the construct number is noted, and on the right, the numbers of expressers and the numbers of transgenic mouse embryos are listed for each construct. **(B-F)** Representative transgenic embryos carrying the constructs described above. The construct number used in each of the transgenic embryos is noted in the right bottom of each figure. ov, otic vesicle; r3, rhombomere 3; r5, rhombomere 5. A, *ApaI*; Sp, *SpeI*; del, deletion; TG, transgenic embryos. Transgenic experiments with construct #5-8 were performed by Dr. Studer.

I then linked the repressor region to the *Hoxb2* r3/5 enhancer and only 22% of the transgenic embryos (n=9) exhibit r3/5 enhancer activity (construct #6). The RARE was deleted in construct #7, which attenuated the repressor activity; all of the embryos showed reporter staining in r3/5 (n=13). I then deleted the Krox20 binding site in the next construct (#8), which also led to an attenuation of the repressor activity. In construct #9, I changed the fifth nucleotide in the Krox20 binding side from a G to a C. This change has been shown to specifically abolish Krox20 binding activity. The majority of transgenic embryos carrying this construct exhibit r3/5 reporter staining (Figure 6.3F). This result shows that the region containing the Krox20 binding site has repressor activity.

6.2.4 Deletion of the Krox20 binding site leads to expansion of reporter expression in r3/5

In order to show that this Krox20 binding site has an important role in restricting *Hoxb1* expression *in vivo,* I used a *Hoxb1* fragment in which a *lacZ* reporter gene had been inserted in frame into the first exon of *Hoxb1* (see also Marshall et al., 1994). Transgenic embryos carrying this construct show restricted r4 reporter staining in the hindbrain. I then tested the same construct in which the Krox20 binding site had

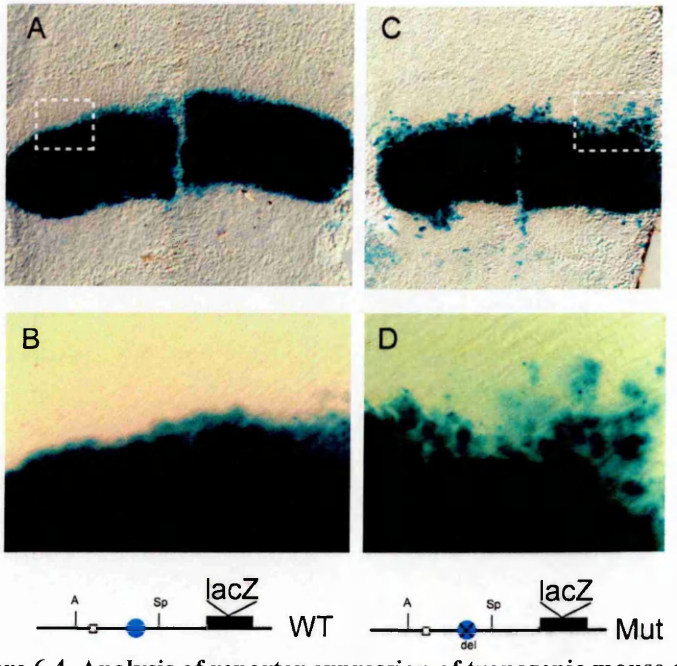


Figure 6.4. Analysis of reporter expression of transgenic mouse embryos carrying an r4 *Hoxb1-lacZ* genomic fragment and a variant in which the Krox20 binding site was deleted.
The figures show flat mounts of the hindbrain of 8.5dpc transgenic mouse embryos. (**A**) Reporter staining in r4 of embryo carrying a WT transgene, containing a *Hoxb1* genomic fragment, in which a *lacZ* reporter gene has been inserted in frame into the first exon of *Hoxb1*. The white dotted box shows the region which is enlarged in figure **B**. **(C)** Reporter staining in r4 of a transgenic mouse embryo carrying the same construct as described for Figure A, except that the Krox20 binding site has been deleted. The white dotted box encompasses the region which is shown enlarged in Panel **D**. The schematic diagrams below the figures illustrate the construct design.

been deleted. The reporter staining of this transgenic mouse is remarkably less restricted in the hindbrain. As shown in Figures 6.3C and D, several cells in rhombomeres r3 and r5 are *lacZ* positive. This implies that this element has an important role in restricting *Hoxb1* expression in r4.

6.3 Discussion

Here I show that Krox20 directly downregulates *Hoxb1* expression in the hindbrain. I identified a highly conserved region which is able to bind Krox20 *in vitro* and I showed that this region has repressor activity. The removal of this region from a transgene construct led to the expansion of reporter expression at the boundaries of r3 and r5.

6.3.1 Krox20 is directly restricting *Hoxb1* expression

The data presented here define an additional mechanism limiting *Hoxb1* expression to r4. *Hoxb1* expression is upregulated in a broad domain within the hindbrain during early development. During the course of hindbrain segmentation its expression becomes more restricted to r4. Retinoids and *vhnf1* have been shown to be involved in this process (Studer et al., 1994; Wiellette and Sive, 2003). In this study I show that Krox20 is able, as an additional factor, to directly restrict *Hoxb1* expression, probably through the interaction with its co-repressors Nab-1 and Nab-2. This observation is consistent with Krox20 overexpression experiments in the hindbrain, which leads to downregulation of *Hoxb1* expression in the hindbrain (Giudicelli et al., 2001).

Both *Nab-1* and *Nab-2* have been shown to be expressed in r3 and r5, and *Nab1* shows upregulation at the r5 boundaries (Mechta-Grigoriou et al., 2000). It has been proposed that Nab proteins regulate the precise transcription level of Krox20, by a negative feedback loop (Mechta-Grigoriou et al., 2000). Overexpression of *Nab-1* and *Nab-2* in zebrafish embryos led to loss of or decreases in the *Krox20* expression domains and the *Hoxb1* expressing territory in r4 was extended more caudally (Mechta-Grigoriou et al., 2000). *Krox20* has been also shown to repress expression of *follistatin* in r3 but not in r5 (Seitanidou et al., 1997). Though this in-

teraction between these two genes was only shown genetically, it is likely that the repression mechanism is direct with the involvement of the co-repressors Nab1 and Nab2.

In zebrafish, the Krox20 binding site was only present in the regulatory region of the co-paralogous gene *Hoxb1a*, not in *Hoxb1b*. *Hoxb1b* is only transiently expressed in r4 and is then gradually downregulated, whereas *Hoxb1a* is upregulated at later stages and maintained during hindbrain development by an autoregulatory mechanism (McClintock et al., 2002; Prince and Pickett, 2002). Therefore, only the stable expression of *Hoxb1a* is restricted by Krox20, whereas restriction of the transiently early expression of *Hoxb1b* appears not to occur.

6.3.2 The regulation of Krox20 and Hoxb1 Expression is tightly linked

The described mechanism ensures the separation of cells in even- and in odd-numbered rhombomeres from each other. *Krox20*, which is expressed in r3/5, defines the odd-numbered cells. Krox20 null mutants have been generated by targeted inactivation and mutant phenotype shows losses of r3 and r5 (Schneider-Maunoury et al., 1993). Therefore Krox20 is important to maintain the odd-numbered rhombomeres, r3 and r5. Furthermore, it is involved in determination of segmental identity in r3 and r5 by regulation directly the expression of several *Hox* genes including *Hoxa2, Hoxb2* and *Hoxb3* (Manzanares et al., 2002; Nonchev et al., 1996b; Sham et al., 1993). It also plays an important role in restricting the intermingling of cells between odd and even-numbered rhombomeres by directly regulating a member of the *Eph* family, *EphA4* (Theil et al., 1998).

Hoxb1 on the other hand has been shown to play a critical role in establishing the r4 identity. In *Hoxb1* homozygote mutants, the r4 is transformed to an r2 identity (Studer et al., 1996).

A regulatory link between Krox20 and Hoxb1 has been observed previously. In the absence of Hoxa1 and Hoxb1, *Krox20* fails to be expressed in r3 (Barrow et al., 2000; Rossel and Capecchi, 1999). This suggests that Hoxb1, directly or indirectly, ensures the expression of Krox20 in r3, and thereby establishing the adjacent anterior rhombomere of r4 as an odd-

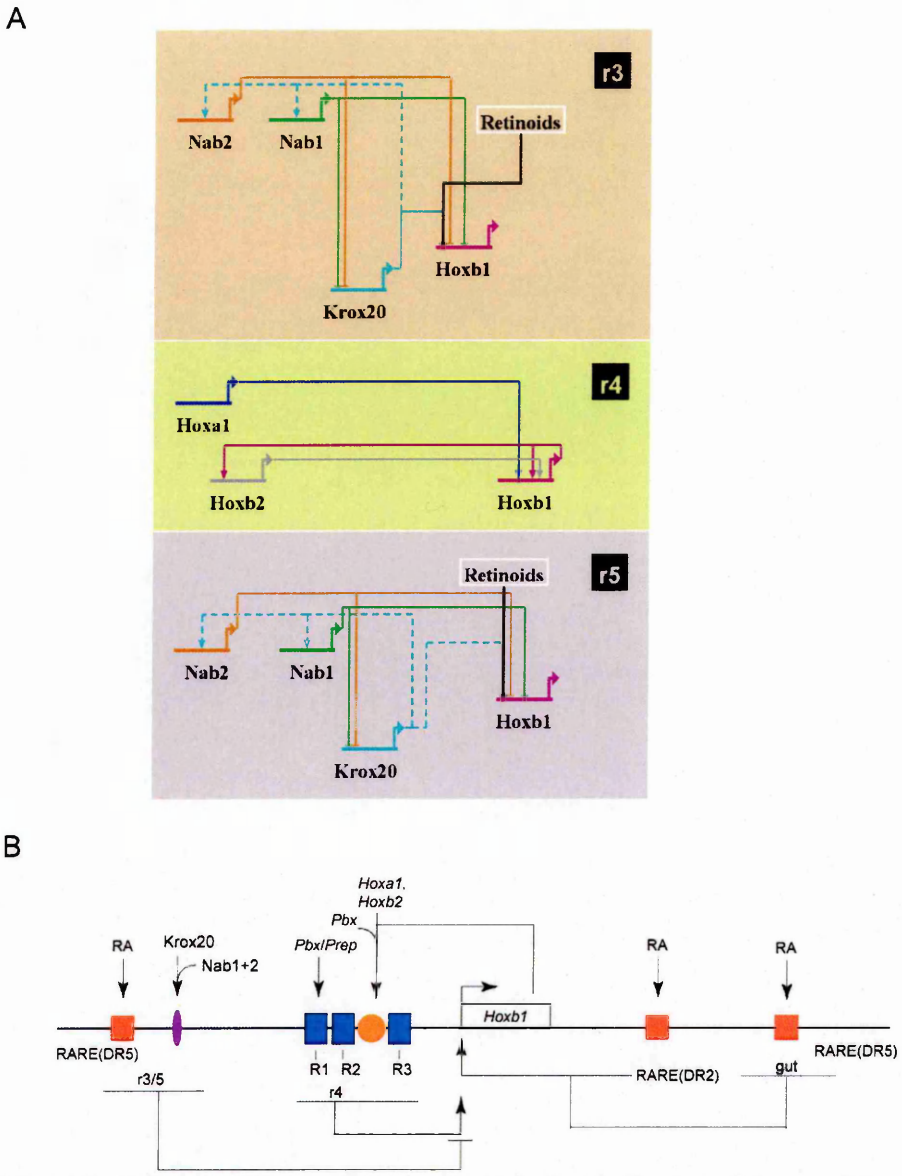


Figure 6.5. Schematic diagrams of the regulatory network in the hindbrain and the regulatory module of *Hoxb1*.
(A) Genetic regulatory network in r3, r4 and r5. Dotted line illustrates that the interaction between the genes have been not shown to be direct. Arrowhead at the end of the line represents activation, whereas perpendicular line at the end of the line illustrates repression between the connected genes. **(B)** The regulatory module of *Hoxb1*. On the top of the module the *trans*-acting factors are listed, on the bottom of the module the embryonic expression domains for each of the modules are listed.

numbered identity.

The regulation of *Hoxb1* and *Krox20* is tightly linked which ensures the maintenance and the establishment of the identity of r3, r4 and r5. My work showed that Krox20 has a direct effect on the regulation of *Hoxb1* expression.

Chapter 7

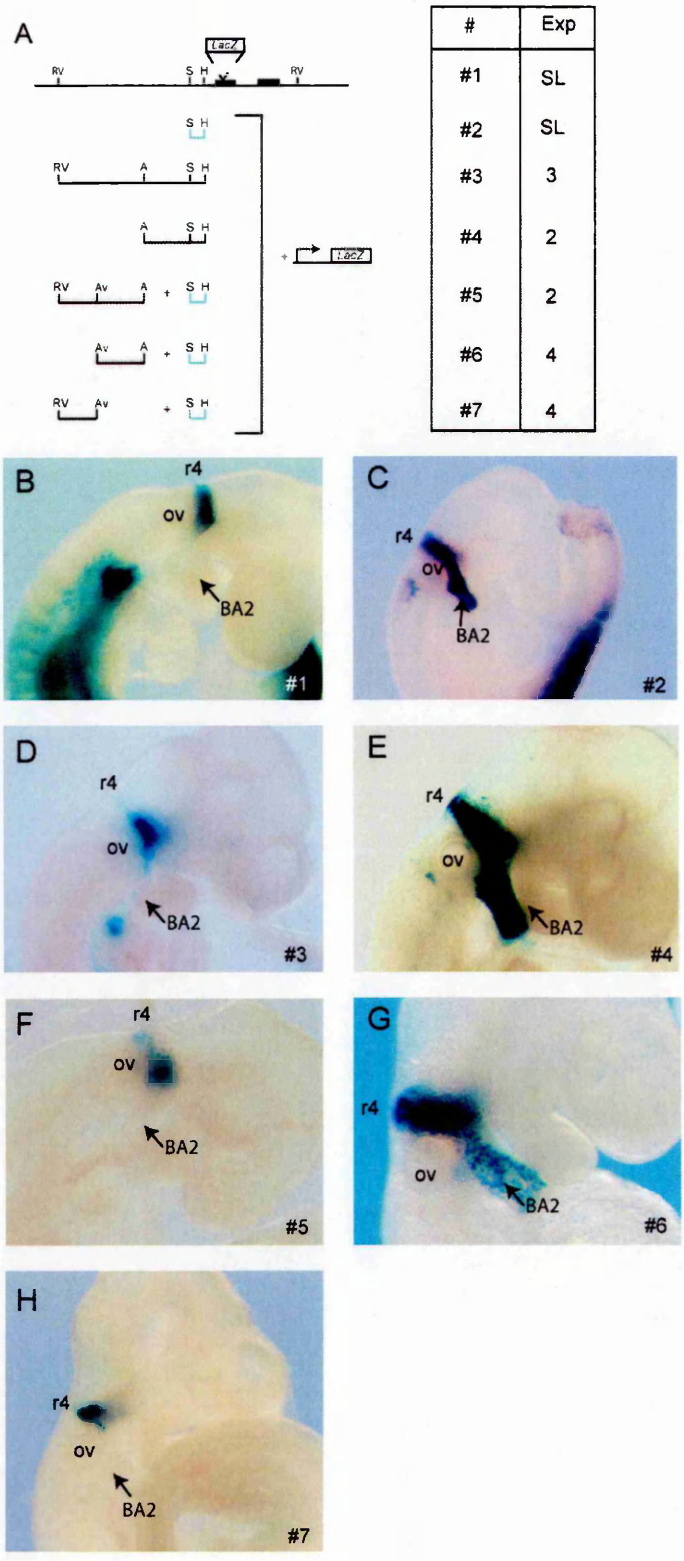
A conserved element represses Hoxb1 expression in mesenchymal neural crest cells

7.1 Introduction

In mouse and chick embryos, *Hoxb1* is expressed in various embryonic tissues, including rhombomere 4, the facial ganglion (a neurogenic crest derivative), the lateral mesoderm, the primitive streak, the node and the somites. In addition, *Hoxb1* is expressed early in all the neural crest derivatives migrating from r4 into the second branchial arch (BA), but it becomes rapidly down regulated in mouse embryos and remains on only in the neurogenic component, the facial ganglion. In contrast, in chick embryos, *Hoxb1* expression stays on during development and is not downregulated.

Several studies have shown that *Hoxb1* expression becomes restricted during development (Fox, 2000; Samad et al., 2004; Studer et al., 1994). A direct repression mechanism has been shown for r3/5 repression mediated by retinoids (Studer et al., 1994). However, the basis for *Hoxb1* downregulation in neural crest is unknown.

Here, I will present data showing that the *Hoxb1* NCC repressor is located in the intergenic region of *Hoxb1/b2,* is cable of repressing the rhombomere-specific enhancer activities of *Hoxa2* and *Hoxb2*, and is highly conserved among different mammals. Further, I will present the method used for removing the *Hoxb1* repressor from its endogenous locus and discuss the potential for further analysis of the animal harboring the mutation.



#	Exp
#1	SL
#2	SL
#3	3
#4	2
#5	2
#6	4
#7	4

Figure 7.1. Mapping the *Hoxb1* BA repressor.
(A) Diagram of various constructs tested for generation of transgenic embryos. In construct #1, the *lacZ* gene was fused in frame with the first exon of *Hoxb1* (Marshall et al., 1994). The other constructs were all linked to the *lacZ* reporter gene. Blue region shows the *StuI-HindIII* fragment which contains the *Hoxb1* r4 enhancer. On the right, the table contains the construct number (#) and the number of embryos obtained showing the same expression pattern with a given construct. SL (stable line) indicates that a stable line was analyzed for this construct. **(B-C)** Lateral views of analyzed embryos. The numbers of each construct used for each transgenic embryo are indicated on the right lower corner of each figure. BA2, 2nd branchial arch; ov, *otic vesicle*; r, rhombomere: R, *EcoRV*; A, *ApaI*: S, *SpeI*; H, *HindIII*.

7.2 Results

7.2.1 Identifying the NCC repressor in the intergenic region of *Hoxb1/2*

I analyzed the reporter expression in the second BA in two stable lines over time (Figure 7.1). The first line carries a 20kb genomic fragment containing the Hoxb1 locus in which a *lacZ* reporter gene has been inserted in frame into the first exon of Hoxb1 (construct #1). The expression of the reporter is initially strong in the second BA, but it is rapidly downregulated at around 9.5dpc (Figure 7.1B and data not shown). In the second line, which carries only the r4 enhancer of *Hoxb1* (*StuI-HindIII* fragment), the reporter expression in the second BA was strongly upregulated throughout early development (Figure 7.1C and data not shown). This suggests that the repressor is located either up- or downstream of the r4 enhancer, so I used deletion analysis to map its position.

I linked a 4.7kp *EcoRV-HindIII* fragment, containing the r4 enhancer and upstream sequence, to the *lacZ* reporter gene (construct #3). Transient transgenics were generated at analyzed at 9.75dpc. Embryos carrying this transgene exhibited no reporter gene expression in the BAs (Figure 7.1D). This suggests that the repressor region is located in this 4.7kb fragment. I subdivided and tested various fragments in combination with the r4 enhancer.

I first analyzed a 1.7kb *ApaI-HindIII* fragment in this reporter gene assay (construct #4). Transgenic mouse embryos showed strong reporter expression in the second BA (Figure 7.1E) indicating the repressor was not present in this segment. Therefore I tested the more upstream sequence in a *lacZ* reporter construct (#4) containing the 2.8kb fragment linked to the r4 enhancer. This construct produced transgenic embryos with high levels of reporter staining in r4 but no reporter expression in the second BA, suggesting that the repressor is located in this 2.8kb fragment (Figure 7.1F). I then cloned two fragments, one containing a 1.5kp *AvrII-ApaI* region of the 2.8kb fragment (construct #6) and the other one a 1.3kp *EcoRV-AvrII* region (construct #7). These constructs were then linked to the *Hoxb1* r4 enhancer. I observed

reporter gene expression of the first construct in the second BA (Figure 7.1G), whereas the second construct gave no expression in the second BA (Figure 7.1H). This locates the repressor in the 1.3kb *EcoRV-AvrII* fragment.

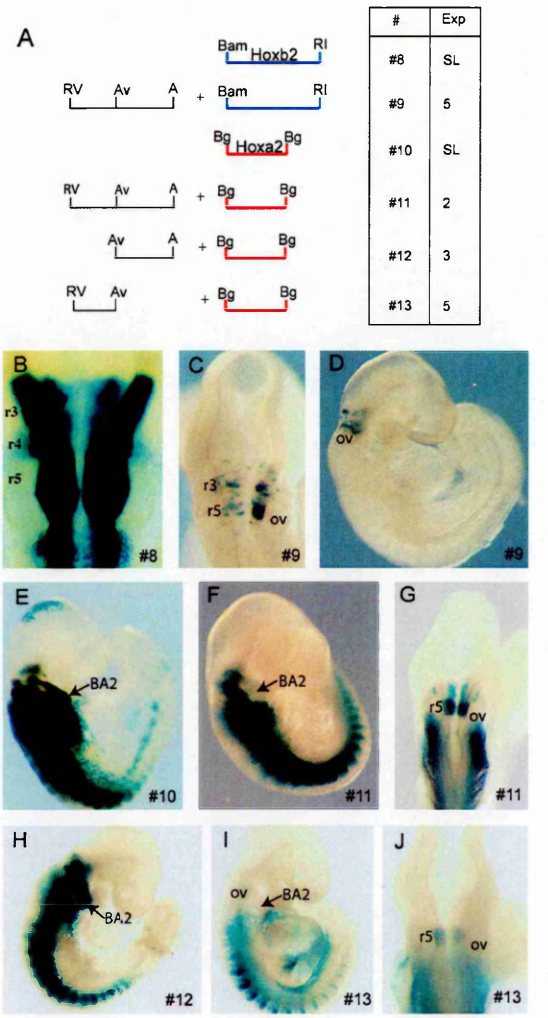


Figure 7.2. Testing the *Hoxb1* constructs on heterologous enhancers of *Hoxb2* and *Hoxa2*.
(A) Diagram of the construct used for analyzing the *Hoxb1* repressor. For constructs #8 and #9, the r3/5 and r4 elements of *Hoxb2 (BamHI-EcoR* fragment) were used (Maconochie et al., 1997; Sham et al., 1993). For constructs #10 through #13 the *Hoxa2 BglII* fragment was used to direct r3/5 reporter expression (Nonchev et al., 1996b). On the right, the table contains the construct number (#), the number of transgenic mice obtained (TG), and the number of embryos obtained that show the same expression pattern with a given construct. SL (stable line) indicates that a stable line was analyzed for this construct. **(B-J)** Lateral and dorsal views of the analyzed embryos carrying the listed transgenes. The construct numbers used for the transgenic embryo are indicated on the right lower corner of each figure. BA2, 2nd branchial arch; ov, *otic vesicle*; r, rhombomere; RV, *EcoRV*; A, *ApaI*; Av, *AvrII;* Bam, *BamHI*; RI, *EcoRI*, Bg, *BglII.*

7.2.2 Testing the Hoxb1 repressor on heterologous enhancers

Next I assayed the *Hoxb1* repressor activity on heterologous enhancers of *Hoxb2* and *Hoxa2*. The *BamHI-EcoRI* fragment of *Hoxb2* directs expression in r3/5 and r4 (construct #8) (Figure 7.2A and see Maconochie et al., 1997). I linked the *Hoxb1 EcoRV-ApaI* fragment to

the *Hoxb2* enhancer and observed strong downregulation of reporter expression in the hindbrain (Figures 7.2 C, D).

The *BglII* fragment of the *Hoxa2* enhancer controls expression in r3/5 and in NCC cells (#10 and see Nonchev et al., 1996a; Nonchev et al., 1996b). When the *EcoRV-ApaI* fragment was linked to this reporter construct (construct #11), I observed downregulation of reporter staining in r3 and in the second BA (Figures 7.2 F, G).

I then divided the fragment in half. The *Hoxb1 AvrII-ApaI* fragment was unable to downregulate the Hoxa2 enhancer (construct #12 in Figure 7.2H) while the *EcoRV-AvrII* fragment, when linked to the *Hoxa2* enhancer, resulted in dramatic downregulation (construct #13). Comparison of the repressor activity of the *EcoRV-ApaI* to that of the smaller *EcoRV-AvrII* fragment, suggests that the activity of the repressor element is distance-dependent.

7.2.3 Alignment of the repressor regions reveals conserved regions

I performed a MACAW alignment of the *Hoxb1/2* intergenic regions of different species (Figure 7.3A) and identified two conserved regions (CRI and II) in the *EcoRV-AvrII* fragment and one conserved region (CRIII) in the *AvrII-ApaI* region (Figure 7.3A). This conservation is limited to mammals; I was unable to identify any conservation in chick or fishes in these regions (Figures 7.3.A, B). I cloned these regions and assayed their repressor activity with a *Hoxa2* enhancer reporter construct. CRI has strong repressor activity and as shown in figure 3C, almost the entire reporter expression is absent, leaving only weak reporter staining in r5. I also tested the repressor activity of the CRI on the Hoxb1 enhancer (construct #15) (Figures 7.3E, F). At 9.25dpc, I did not observe any downregulation of reporter staining in the second BA (Figure 7.3E). However, I observed downregulation of reporter expression at later stages, such as 10.25dpc (Figure 7.3F). This is consistent with the earlier observation that reporter expression is strong at around 9.5dpc but is rapidly downregulated at later stages. Further I tested the repressor activates of CRII and III, but these constructs did not exhibit any downregulation (Figures 7.3G, H). These findings suggest that region CRI exclusively medi-

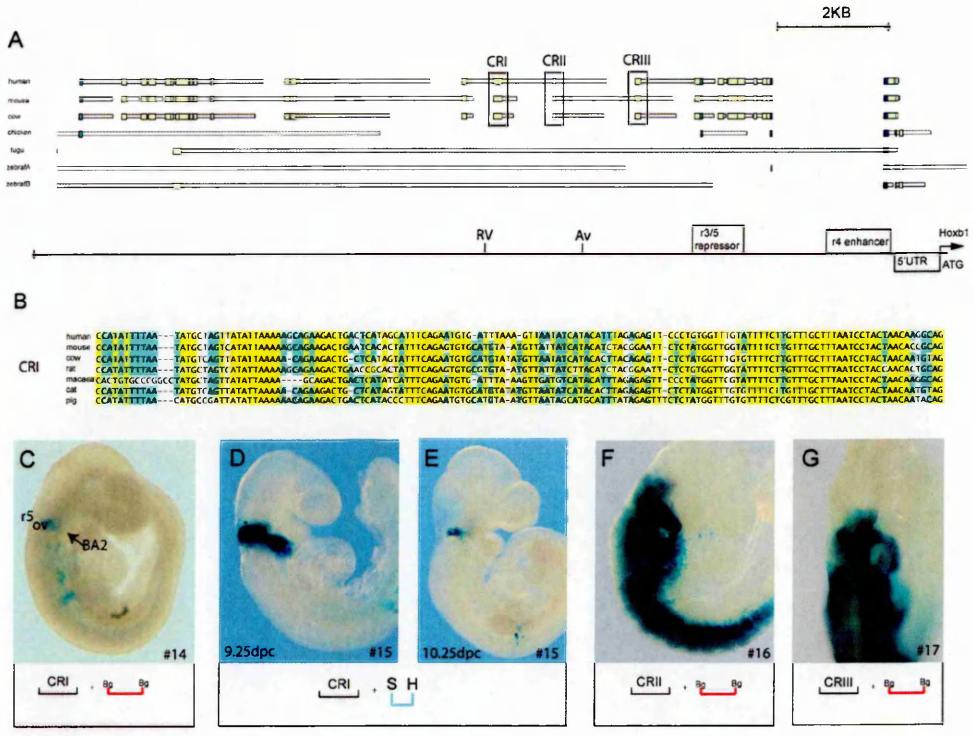


Figure 7.3. Alignment of the *Hoxb1/b2* intergenic region and analysis of conserved regions using transgenic embryos.
(A) MACAW alignment of the *Hoxb1/b2* intergenic region using various species, including human, mouse, cow, chicken, fugu, and zebrafish(a) and (b). Location of known elements and restrictions sites (RV, *EcoRV* and Av, *AvrII*) are illustrated below. Boxes in the alignment encompass conserved regions (CRI-III) analyzed here. **(B)** Local sequence alignment of the first conserved region (CRI). **(C-F)** Representative embryos transgenic for construct shown below each figure. Note that, embryos **D** and **E** were analyzed at different developmental stages.

ates the BA specific downregulation of *Hoxb1* expression.

In order to evaluate the function of this element, I decided to remove it from its endogenous locus (Figure 7.4). I designed a targeting construct, in which the CRI region was replaced by a Neomycin resistance cassette (Figure 7.4A). The ES cells went germ line and further analysis has to be performed to evaluate the phenotype.

7.3 Discussion

Hoxb1 shows early expression in various tissues but becomes downregulated during the course of development. In the last Chapter (Chapter 6), I showed that Krox20 directly downregulates *Hoxb1* expression in r3 and r5. In this study, I investigated the nature of the downregulation of *Hoxb1* expression in mesenchymal NCC migrating into the second BA. I showed that the repressor region is located in the intergenic region of *Hoxb1/b2*, and it is

highly conserved among different mammals. Further I showed that this element is able to repress heterologous enhancer activity. Finally, I present the strategy I used to remove the conserved region from its endogenous locus.

7.3.1 *Hoxb1* is downregulated in the mesenchymal components of NCC in mouse embryos

Cranial NCC derived from the dorsal hindbrain give rise to mesenchymal and neurogenic components. They migrate in discrete streams into the first three BA. *Hoxb1* and *Hoxb2* are only expressed transiently in the mesenchymal components of the second BA in mouse embryos. This coupled downregulation is due to the fact that Hoxb1 directly regulates *Hoxb2* expression (Maconochie et al., 1997).

Hoxb1 expression is only maintained in the neurogenic component of the NCC migrating into the second BA, which will give rise to motor neurons of the facial cranial nerve. In *Hoxb1*-null mice, motor neurons were generated in r4, but they were incorrectly specified (Goddard et al., 1996; Studer et al., 1996). They fail to migrate to their correct position, and they did not project to the proper peripheral targets. Instead, these motor neurons acquired characteristics of r2 trigeminal motor neurons (Studer et al., 1996).

In contrast, in mouse embryos, *Hoxa2* is the only *Hox* gene that is continuously expressed in mesenchymal cells in the second BA, in which it plays a conserved role as a selector gene for the derivatives of the NCC (Gendron-Maguire et al., 1993; Rijli et al., 1993).

7.3.2 Differences in expression and function of Hoxb2 in different species

The downregulation of *Hoxb1* expression in the mesenchymal components appears to be limited to mammals. This finding is reflected in the functional differences of the Hoxb1 target gene, *Hoxb2*, between mouse and zebrafish. In mouse embryos, *Hoxb2* is downregulated early in the second BA, whereas, in zebrafish, it stays on in the mesenchymal NCC during

later stages (Hunter and Prince, 2002). This is consistent with my regulatory studies on the *Hoxb1* mesenchymal repressor region; I was only able to identify the CRI region in mammalians, whereas, in chick or in fishes, this region appears to be absent.

Hoxb2 mutant mice do not exhibit any defects in the derivatives of the mesenchymal

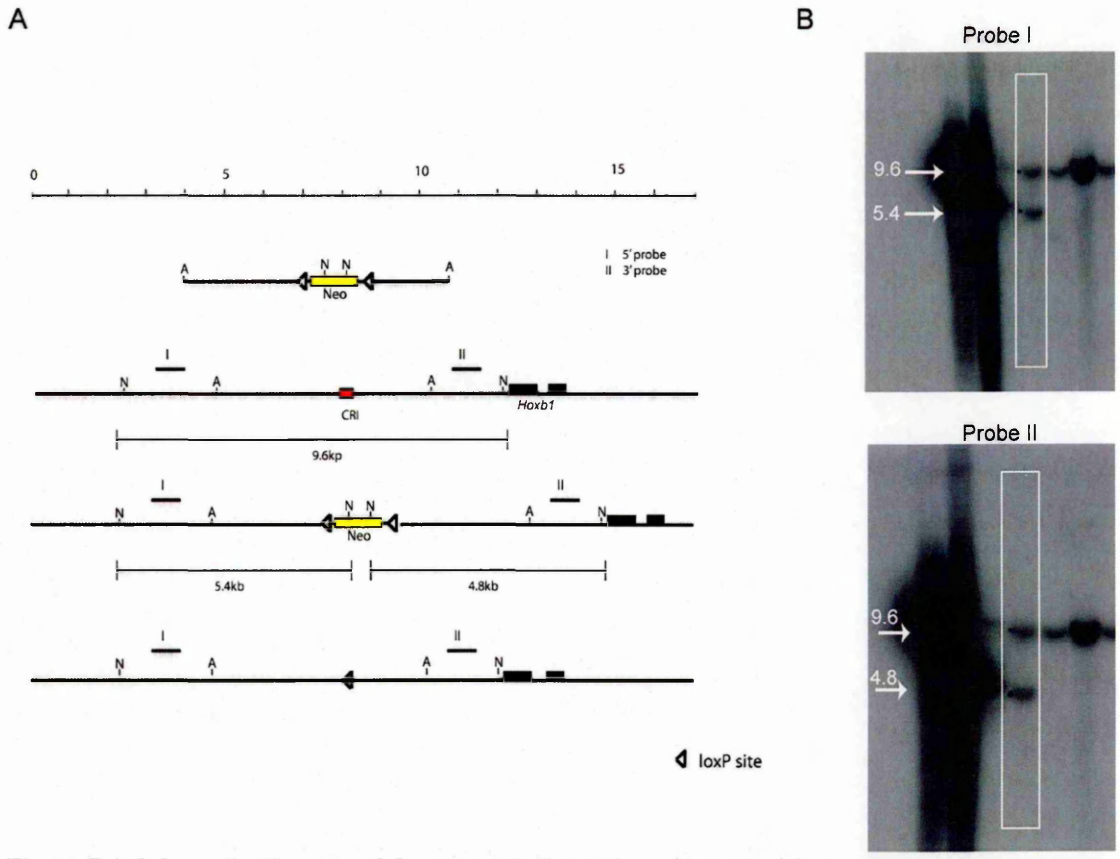


Figure 7.4. Schematic diagram of the targeted disruption of the *Hoxb1* repressor region.
(A) Diagram showing the structure of the wild type *Hoxb1* locus, including the location of the Hoxb1 coding sequence, and of the CRI position, the targeting construct, the targeted locus after successful recombination and the targeted locus after the removal of the *loxP* flanked Neomycin resistance cassette by cre driver. I and II are the regions of the probes used for the Southern gel transfer. N, *NotI*; A, *ApaI*; Neo, Neomycin resistance gene. **(B)** Embryonic stem cell DNA blot hybridized with probe I (top) and probe II (bottom) after digestion of the DNA with *NcoI*. The numbers illustrate the sizes of the bands in kb. N, *NotI*; A, *ApaI*, Neo, Neomycin resistance cassette; CRI, conserved region I.

components of NCC, however, the development of the facial nerve was affected (Barrow and Capecchi, 1996; Davenne et al., 1999). In contrast, in zebrafish, *Hoxb2* has been shown to function as a selector gene; ectopic expression of *Hoxb2* in zebrafish embryos results in transformation of the first arch structures into a second arch identity (Hunter and Prince, 2002). However, in zebrafish, both *Hoxa2* and *Hoxb2*, have to be knocked-down to achieve a mirror-image duplication of first arch structures (Hunter and Prince, 2002). In mouse embryos, the removal of *Hoxa2* is sufficient to cause a transformation of the derivatives of the second BA

(Gendron-Maguire et al., 1993; Rijli et al., 1993).

This suggests that *Hoxb1* expression in zebrafish stays on in mesenchymal NCC in order to maintain *Hoxb2* expression, which, in conjunction with *Hoxa2,* patterns the second BA. In mouse, Hoxa2 itself is sufficient for patterning the second BA (Gavalas et al., 1997; Gendron-Maguire et al., 1993; Rijli et al., 1993).

7.3.3 Generating of a mutant in which the CRI element is removed

I successfully generated a mutant mouse in which the CRI region is removed from its endogenous locus (Figure 7.4). This mutant will provide information about the function of the repressor region, CRI, in mammals. Furthermore, analysis of the mutant may help to explain why *Hoxb1* expression in the mesenchymal components of the second BA is downregulated only in mammals and not in chicken or fish.

It will also be interesting to determine whether the CRI deletion mutant can rescue the Hoxa2 knock-out second BA transformation. In the repressor deletion mutant, Hoxb1 expression may stay on in the second BA, and this could result in maintenance of *Hoxb2* expression in this domain. *Hoxb2* is a paralogous gene of *Hoxa2* and may functionally compensate for Hoxa2 to pattern the second BA structures. For example, in zebrafish, it has been shown that ectopic expression of *Hoxb2* causes the first BA elements to be transformed to a second BA identity (Hunter and Prince, 2002).

To further investigate the basis of the downregulation of *Hoxb1* expression in the mesenchymal components of the second BA, it will be important to understand which *trans*-acting factor(s) mediates the downregulation of *Hoxb1* expression in BA2. This study will require biochemical analysis or a candidate gene approach.

Chapter 8

Collaborative work

8.1 The distal enhancer contains a conserved Hox/Pbx site required for appropriate expression of the RARB locus in the hindbrain.

8.1.1 Introduction

Studies of my collaborators show that *RARβ* uses a two-step transcriptional regulatory mechanism. The early neural expression is regulated by retinoic acid (RA), which is produced by the mesoderm. At later stages (10.5dpc) the expression of *RARβ* becomes positively regulated by overlapping inputs from Hoxb4 and Hoxd4. This late *Hox*-dependent expression is initiated at the time of morphological segmentation, serving to sharpen and fix the expression border of RARβ at the r6/7 boundary.

The *RARβ* gene consists of two promoters, which are regulated by two different enhancers; a distal enhancer that encompasses the distal (P1) promoter and a proximal enhancer that includes the proximal (P2) promoter (Figure 8.1A). When transgenic mice carrying the 2.3kb *NheI* distal promoter were crossed with double homozygous mice lacking the functions of Hoxb4 and Hoxd4, the enhancer activity was abolished. Therefore the distal enhancer activity requires Hoxb4 and Hoxd4 function. But it is unknown whether the regulation of the distal enhancer is directly or indirectly regulated by Hoxb4 and Hoxd4.

8.1.2 Results

To investigate whether the regulation of distal enhancer activity by *Hox4* genes might be direct, the sequences contained in the 2.3 kb *NheI* fragment containing the distal *RARβ* promoter was examined for potential Hox responsive elements. A putative bipartite consensus sequence for interaction with Hox and Pbx proteins (PH) was identified about 75 bases upstream from the characterized *RARβ1/3* transcription start site (Mendelsohn et al.,

1994a). This element is conserved in syntentic regions upstream of the *RARβ* locus in other mammalian genomes, including human, chimp, rat and dog (Figure 8.1B).

To determine whether the Hoxb4 protein was capable of binding to the putative Hox/Pbx motif, I used a double-stranded oligonucleotide spanning the PH element (Figure 8.1B) in electrophoretic mobility shift assays (EMSA) (Figures 8.1C, D). In the EMSAs, the PH site in *RARβ* displayed concentration dependent binding of Hoxb4 protein in a manner similar to that seen using a previously characterized Hoxb4 binding site (HS1+HS2; from the Hoxb4 gene itself as a control (Figure 8.1C). Moreover, in both cases the bound form of oligonucleotides could be shifted to a larger molecular weight complex, by the addition of antibody directed against Hoxb4 (Figure 8.1D). Finally, in EMSA competition assays, the oligonucleotides corresponding to the PH and the HS1+HS2 sites effectively compete with one another for Hoxb4 binding (Figure 8.1D). However, mutated variants of these sequences are unable to compete effectively for Hoxb4 binding, verifying the specificity and nature of the interaction (Figure 8.1D).

In order to test whether the conserved PH site sequences are functionally relevant *in vivo*, their activity was monitored with a *lacZ* reporter gene by in ovo electroporation into developing chick embryos. As observed in transgenic mice, the 2.3 kb *NheI* fragment containing the PH element efficiently mediated reporter expression in the up to the r6/7 boundary in the chick hindbrain (Figure 8.1E). Furthermore, when an oligonucleotide spanning the PH element was multimerized to form three tandem copies, it functioned as an enhancer directing reporter expression up to the same r6/7 boundary in the hindbrain (Figure 8.1F). This reveals that the PH site is sufficient to mediate restricted neural expression of the appropriate pattern. Conversely, when the PH site was deleted in the context of the 2.3 *NheI* fragment, reporter expression was completely abolished in the chick hindbrain showing this site is essential for enhancer activity (Figure 8.1G). Taken together, these data support a direct role for *Hox* in the regulation of the *RARβ* gene.

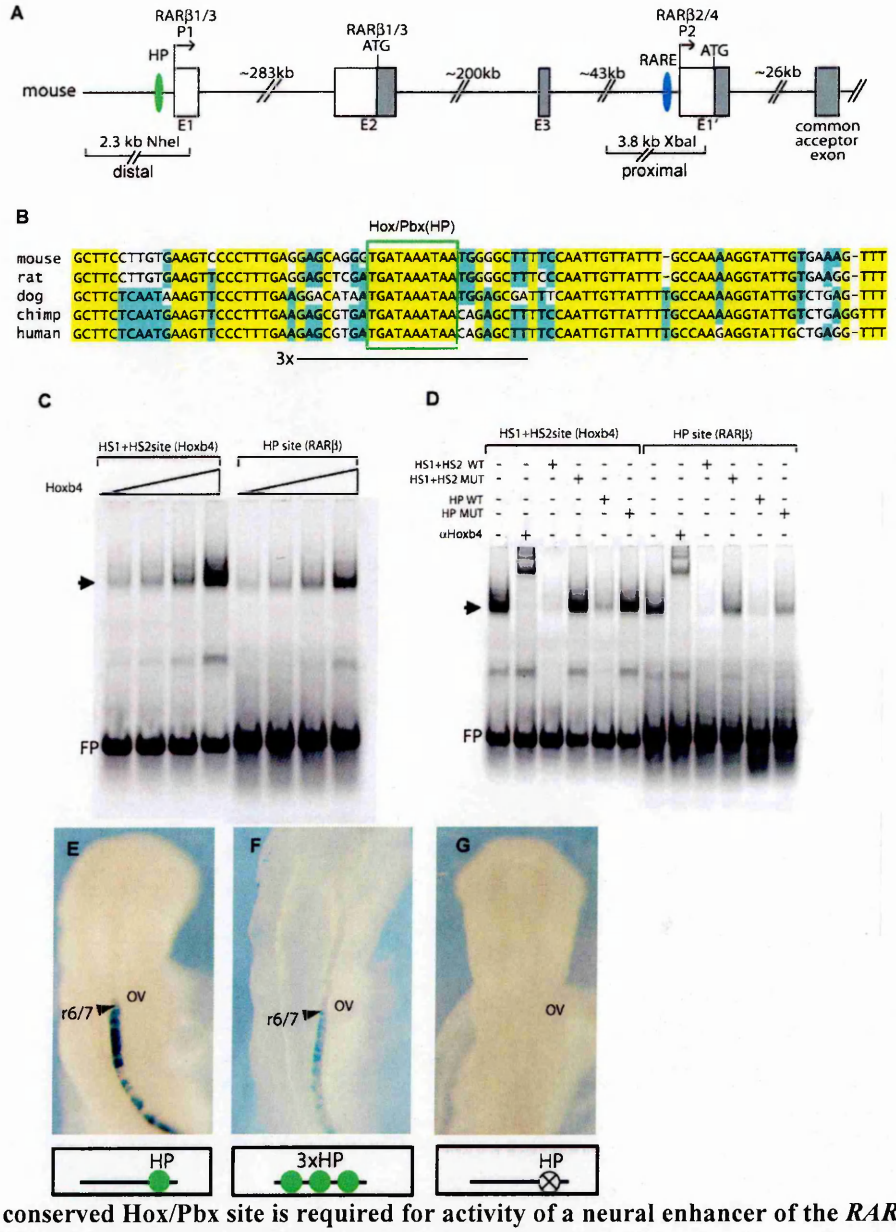


Figure 8.1. A conserved Hox/Pbx site is required for activity of a neural enhancer of the *RARB* locus. **(A)** Map of the 5' end of the *RARB* gene. Transcripts for the Rarβ1 and β3 isoforms initiate transcription from the P1 promoter, while those of the β2 and β4 isoforms initiate at the P2 promoter. The positions of the RARE (grey oval; proximal enhancer), important to the regulation of the β2/4 isoforms (de The et al., 1990; Hoffmann et al., 1990; Sucov et al., 1990) and the newly identified Hox/Pbx element (blue oval, distal enhancer), important for the appropriate regulation of the β1/3 isoforms, are shown. The white boxes and portions of boxes indicate non-coding transcript sequences, grey boxes are the coding sequences. The position of the 2.3 NheI fragment (used in the RARβ1/LacZ transgenic, Mendelsohn et al., 1991) and the 3.8kb XbaI fragment (used in the RARβ2/LacZ transgenic, Mendelsohn et al., 1991) are also shown. Exon E1' is equivalent to E4 in some previous publications, but has been relabelled here to emphasize that it is the start of the new RARβ2/4 transcript and not a contiguous exon of the RARβ1/3 transcript. **(B)** A portion of the MULTIZ alignment (Blanchette et al., 2004) around the Hox/Pbx element is shown. The consensus Hox/Pbx element is boxed in purple. Below the alignment, the murine sequences that were multimerized for the construct used in panel F, are indicated (3x). **(C, D)** Electrophoretic mobility shift assays, where a labelled double-stranded oligonucleotide containing the HS1+HS2 site *(Hoxb4)* (Gould et al., 1997) and the newly identified PH site *(RARβ)* have been incubated with **(C)** increasing amounts of Hoxb4 protein or **(D)** with a constant amount of Hoxb4 protein in combination with Hoxb4 antibody (αHoxb4) or cold competitor oligonucleotides, as indicated. **(E, F, G)** *LacZ* reporter expression in chick embryos electroporated on the right side of the neural tube. The respective constructs are indicated below each panel. The position of the otic vesicle (ov) and the most anterior expression at the r6/7 boundary (closed arrow head) is shown. The blue circle indicates the PH binding site(s). In the construct with the circle with an x, the PH site has been deleted by site directed mutagenesis.

The transgenic analysis presented here demonstrates that the mechanisms regulating both the early and late phases of *RARβ* expression within the hindbrain operate at the transcriptional level. At E8.5, the proximal retinoid-dependent enhancer recapitulates the initial diffuse and transient *RARβ* expression up to the presumptive r5/r6 border. In contrast at E9.5-E10.5, the distal Hox-dependent enhancer maintains the stable segmental border at r6/r7 through the direct inputs from both Hoxb4 and Hoxd4. Together, the summation of the activities from these proximal and distal enhancers accounts for the transcriptional mechanism that governs the establishment and maintenance of the dynamic hindbrain expression of the endogenous *RARβ* gene.

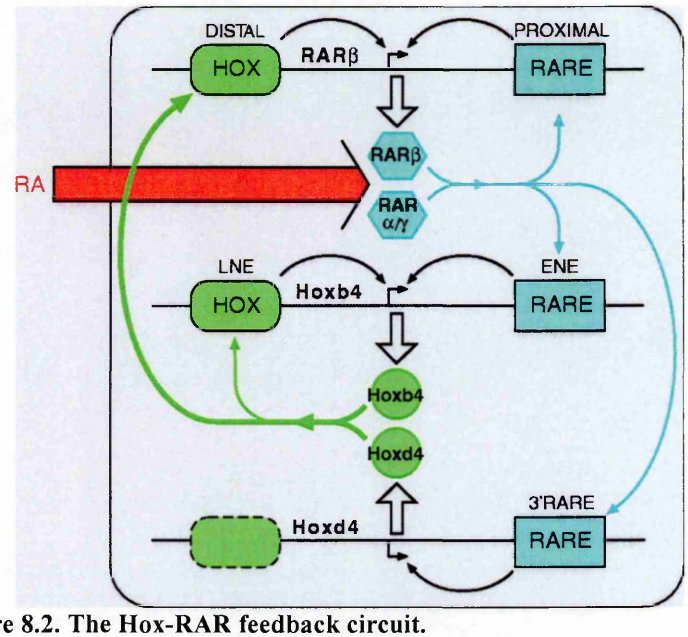


Figure 8.2. The Hox-RAR feedback circuit.
Cartoon shows the *Hox*-responsive enhancers (HOX, green) and RA-responsive enhancers (RARE, blue) from *RARβ* (top), *Hoxb4* (middle) and *Hoxd4* (bottom). Enhancer-promoter interactions (black arrows), transcription and translation (unfilled arrows), and transcriptional regulation by Hoxb4 and Hoxd4 proteins (green arrows) and RARs (blue arrows) are shown. The external RA signal that initiates this neural genetic circuit (red arrow) is synthesized by Raldh2 in the adjacent mesoderm. It is not yet clear whether *Hoxd4* possesses a suitable HOX element (dotted).

8.1.3 Discussion

In this collaborative work, I showed that Hoxb4 directly regulates the late expression of *RARβ*. For this purpose, I used sequence alignment analysis and chick electroporation assay experiments. I was able to identify a highly conserved region within the 2.3kb distal enhancer which is cable of binding Hoxb4 protein. Further I show that a multimerized version of this Hoxb4 binding site linked to a reporter gene is able to direct expression in the neural tube up to the r6/r7 boundary. Further I show that the enhancer activity of this 2.3kb distal enhancer is

abolished, when the Hoxb4 binding site is removed.

Since the homeobox sequences and DNA-binding specificities are very similar between Hoxb4 and Hoxd4, it is likely that Hoxd4 mediates the *RARβ* expression in the same direct manner.

This finding allows us to draw a complex genetic regulatory network which involves the genetic interactions of RARβ, Hoxd4 and Hoxb4 (Figure 8.2). It has been shown that the proximal enhancer directly binds retinoic acid receptors (RAR) and responses to RA in cultured cells (de The et al., 1990; Hoffmann et al., 1990). In this study my collaborators and I show that the distal promoter of *RARβ* is directly regulated by Hoxb4 and possibly by Hoxd4. The same mechanism accounts for Hoxb4 regulation, there is a 3' RA response element (ENE) initiating *Hoxb4* expression, and a 5' regulatory element (LNE) maintaining expression by an autoregulatory mechanism and by Hoxd4. Hoxd4 expression is also initiated by RA, but it is unknown how expression is maintained. This complex transcriptional feedback circuit maintains the expression and establishes the expression borders of multiple *Hox* and *RAR* genes at a single segmental boundary.

8.2 The *Hoxb1* enhancer and control of rhombomere 4 expression: Complex interplay between PREP1-PBX1-HOXB1 binding sites.

8.2.1 Introduction

The autoregulatory enhancer of *Hoxb1* directs its segmental expression in the vertebrate hindbrain. Three conserved repeats (R1, R2, and R3) in the enhancer have been described as Pbx-Hoxb1 (PH) binding sites and a single Prep/Meis (PM) binding site has also been characterized (Figure 8.1 and Ferretti et al., 2000; Pöpperl et al., 1995). My collaborators identified an additional PM site (PM2) downstream to the R3 PH site. I performed sequence alignments of the ARE and showed that the PM2 site is highly conserved among different species. Further I performed deletion experiments using chick electroporation and transgenic

mouse embryo assays, showing that the PM2 site contributes to r4 enhancer activity.

8.2.2 Results

I performed sequence alignment of the ARE region from six vertebrate species using Vector NTI's integrated ClustalW (Thompson et al., 1994) alignment program (Figure 8.3B). The PM2 motif is highly conserved in all cases, with the exception of the zebrafish Hoxb1b gene (Figure 8.3B).

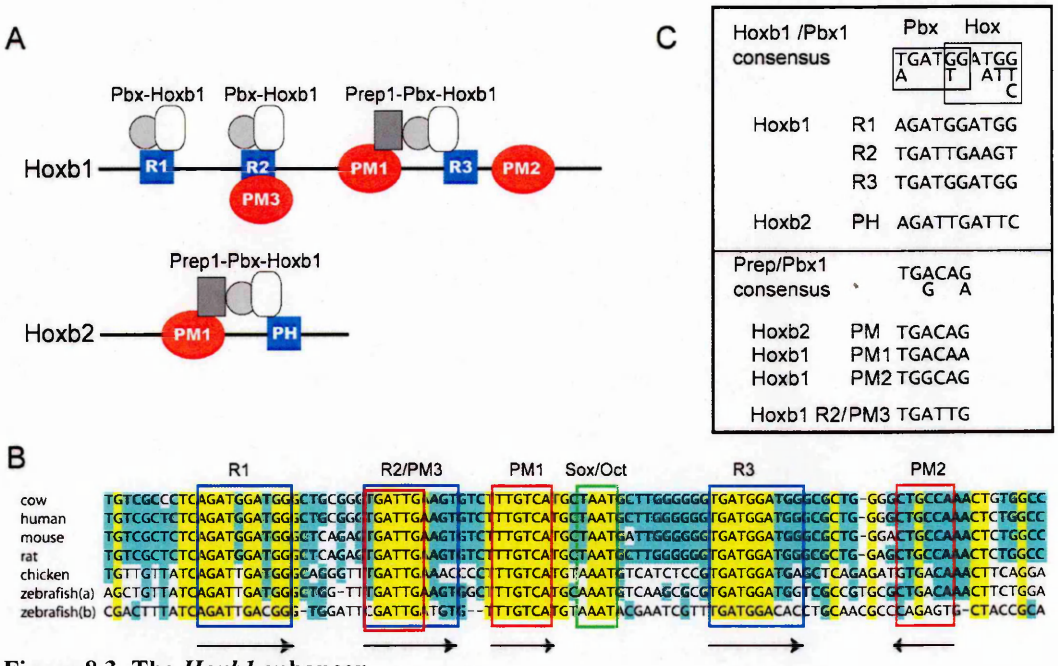


Figure 8.3. The *Hoxb1* enhancer.
(A) Schematic representation of *r4-Hoxb1* and *r4-Hoxb2* enhancers including the Prep/Meis (PM) and Pbx-Hox (PH) sites binding to the various Prep-Pbx complexes. The blue squares indicate the PH sites (R1, R2/PM3, R3 in *Hoxb1* and PH in *Hoxb2*), and the red circles indicate, the PM sites (PM1, PM2). The sequence of the *Hoxb2* PM-PH site is included at the bottom of the *Hoxb2* enhancer scheme. Notice that the R2 site is also called R2/PM3. **(B)** Sequence conservation between mammals, chicken and zebrafish *r4-Hoxb1* regulatory regions. The conserved PH (blue), the PM (red) and the Oct1 (green) sites are boxed. Arrows below the sites indicate site orientation. **(C)** Consensus sequences of the PH and PM sites based upon the mouse *r4-Hoxb1* and *r4-Hoxb2* enhancers.

Therefore, I investigated the relative roles of the PM1 and/or PM2 sites in the context of a *Hoxb1* 622bp fragment containing R1-R3, as well as PM1 and PM2 sites and sequences that serve to restrict expression to r4 (Figure 8.3B). This wild type fragment functions efficiently *in vivo* as an r4 enhancer, as 79% of chick embryos electroporated with this construct display strong staining in r4 (Figure 8.3 and Table 8.1). Mutation of the PM1 site leads to a reduction in efficiency (39% vs. 79%) and many of these embryos display patchy reporter staining further suggestive of a decrease in activity (Figure 8.4B; Table 8.1). Mutation of the

PM2 site in this context had a more pronounced effect on the regulatory activity, as this variant is expressed in only 11% of the embryos, again with a weak, patchy expression only in r4 (Figure 8.4C; Table 8.1). Combining mutations in both PM1 and PM2 sites completely eliminates efficient reporter staining in r4 (Figure 8.4D, Table 8.1). These results show that the

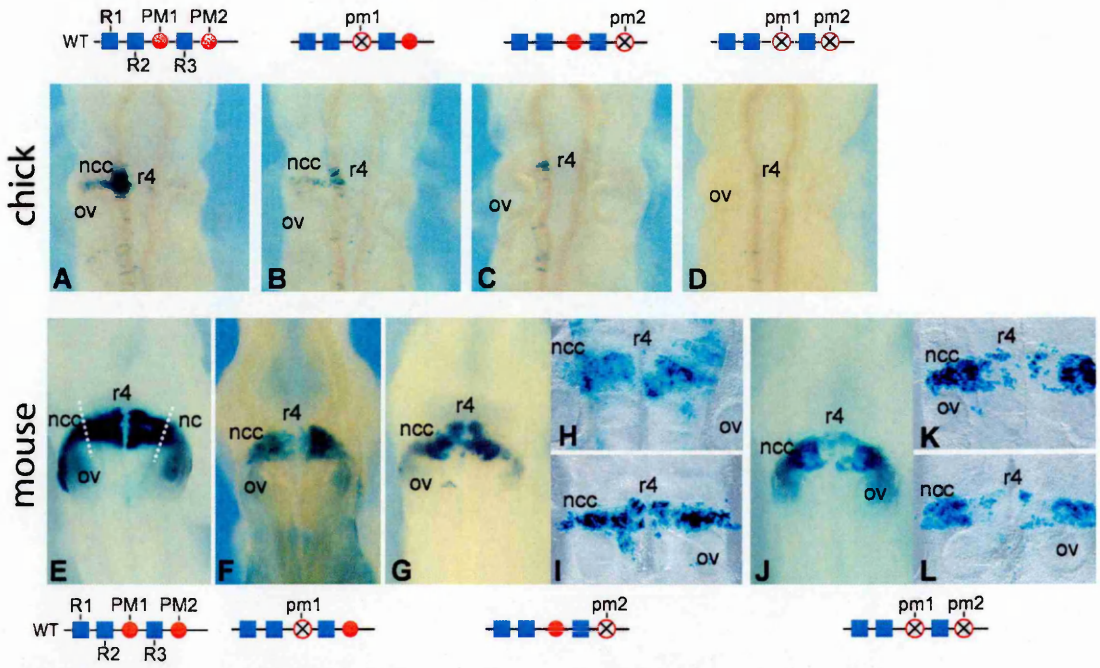


Figure 8.4. Transgenic analysis of the *Hoxb1* enhancer in chicken and mouse embryos.
(A-D) Dorsal views of transgene expression in electroporated chicken hindbrains. They represent expression patterns mediated a wild type 622bp *Hoxb1* fragment **(A)** and variants carrying mutations in the PM1 **(B),** PM2 **(C)** or PM1+PM2 **(D)** sites. The constructs used are noted below each panel. **(E-L)** Represent transgene expression patterns in 9.5 dpc mouse embryos carrying the same wild type **(E)** and mutant PM1 (F), PM2 (G-I) or PM1+PM2 (K-L) sites in the *Hoxb1* constructs. Panels H, I, K, and L show the flat mount preparations of embryos to clearly indicate the patchy and reduced level of expression in r4. The dashed vertical white lines indicate the border between the neural crest cells and the rhombomere or neural tube. Note in J-L that expression is greatly reduced or nearly absent in the r4 territory but is unaffected in the more lateral migrating cranial neural crest cells. This indicates a different requirement for Pm sites in r4 versus neural crest. The arrowhead denotes rhombomere 4; ncc, neural crest cells and ov, otic vesicle.

PM1 and the PM2 sites are required for expression of the reporter in the chick hindbrain in terms of number of positive embryos and relative level of expression in r4.

To further evaluate the *in vivo* regulatory activity of PM1 and PM2, I have tested these same 622bp constructs in transgenic mouse assays, which have been previously employed to evaluate the role of PM1 and R3 in the *Hoxb1* and *Hoxb2* enhancers (Ferretti et al., 2000; Jacobs et al., 1999). In these experiments, the DNA is stably integrated into the chromosome rather than working as an episome as in the chick electroporation experiments. The wild type fragment from *Hoxb1*, containing all the PH and PM motifs, mediated strong reporter staining

in the hindbrain of 9.5dpc mouse embryos (Figure 8.4E), similar to that observed in the chick electroporation experiments (Figure 8.4A). Reporter staining is slightly reduced in the constructs carrying a mutations in the PM1 (Figure 8.4F) or PM2 (Figure 8.4G) sites but the expression remains restricted to r4. Flat-mount preparations of the hindbrains make it easier to see the patchy and reduced expression in r4 of embryos carrying mutations in PM2 (Figures 8.4H, I). Combined mutation of both the PM1 and PM2 sites resulted in a further decrease in expression in the mouse hindbrain (Figure 8.4J). Again flat mounts of hindbrains with the double mutation illustrate that in some cases, expression was abolished in all but a few cells in r4, in particular in the basal plate (Figures 8.4K, L). It is interesting that while there is a loss of expression in r4 and the basal plate in the PM1 and PM1+PM2 mutants, reporter staining is maintained laterally in neural crest cells migrating from r4 (Figures 8.4G-L).

Table 8.1. Summary of chicken electroporation experiments

Construct Name	Number of embryos	Strong r4 expression (percentage)	Weak, patchy or no expression in r4 (percentage)
R1-R2-PM1-R3-PM2	165	79%	21%
m-R1-R2-pm1-R3-PM2	18	39%	61%
m-R1-R2-PM1-R3-pm2	18	11%	89%
m-R1-R2-pm1-R3-pm2	17	0%	100%

Name indicates the structure of the oligonucleotides electroporated (Construct), the number of injected embryos, and the level of reporter gene activity.

8.2.3 Discussion

I showed in this study that the PM2 site of the *Hoxb1* r4 enhancer is highly conserved among different species. This includes cow, human, mouse, rat, chicken, and zebrafish(a) with the exception of zebrafish(b). This is interesting, because the zebrafish *Hoxb1a* gene is expressed in r4 of the developing hindbrain, in a manner similar to the *Hoxb1* gene of other vertebrates, while the duplicated paralogous *Hoxb1b* gene is only expressed transiently in r4 (McClintock et al., 2002). The lack of *Hoxb1b* expression in r4 of the zebrafish hindbrain at later stages was intriguing, because repeats R1-R3 and PM1 are all present and conserved, implying that changes in other motifs must contribute to the absence of segmental expression of

this duplicated gene. Therefore, the diverged sequences in the PM2 site of *Hoxb1b* may explain the differential expression of the duplicated *Hoxb1* genes in zebrafish.

Together these *in vivo* transgenic assays in mouse and chick embryos indicate that the rhombomere-restricted activity of the *Hoxb1* enhancer is dependent upon contributions from both the PM1 and PM2 sites. In the mouse, the role of PM1 and PM2 is similar to that scored in chick embryos. Furthermore, mutation of the PM sites specifically affected transgene expression in r4 cells and not in the r4-derived migrating neural crest cells. This demonstrates the distinct regulatory requirements for neural crest and rhombomeric cell populations.

Chapter 9

General Discussion

In this thesis, I analyzed the regulation of *Hoxa2* and *Hoxb1* during hindbrain development. I identified enhancer elements which direct r2 and r4 expression of *Hoxa2* (Chapters 3 and 4). Further I analyzed the regulatory regions of the two fugu co-paralogous genes *Hoxa2(a)* and *(b)* and identified subtle sequence drift in the rhombomeric elements which explain the differential expression of these two genes (Chapter 5).

I also analyzed the regulatory regions of *Hoxb1* and identified two repressor regions. The first repressor region mediates Krox20-dependent downregulation of *Hoxb1* expression in r3 and r5 (Chapter 6). The second repressor region mediates the downregulation of *Hoxb1* expression in BA2 in mammals (Chapter 7).

For all the projects described in this thesis, I used comparative genomics (phylogenetic footprinting); by comparing orthologous genomic DNA from different species, it is possible as an initial step to identify potential regulatory DNA sequences in conserved regions (Aparicio et al., 1995; Lenhard et al., 2003; Manzanares et al., 2001).

9.1 Comparative genomics as a powerful tool to analyze *cis*-regulatory elements

Recently, the complete DNA sequence of the human and mouse genomes has been determined, the chick genome DNA sequence is projected to be completed soon, and other vertebrate species will follow in the not too distant future (Gibbs et al., 2004; Hillier et al., 2004; Venter et al., 2001). This opens up unparalleled opportunities, not only for exploring gene evolution by comparisons of coding regions, but inter-specific sequence comparisons may also be extremely powerful for identifying regulatory regions controlling the temporal and spatial patterns of gene expression.

The approach of searching for conserved motifs between vertebrate *Hox* genes in combination with biochemistry and functional assays has been very useful in helping to identify

some parts of the *cis*-regulatory machinery controlling the restricted expression of *Hox* genes (Aparicio et al., 1995; Dupé et al., 1997; Frasch et al., 1995; Gavalas et al., 1998; Gould et al., 1998; Huang et al., 1998; Kim et al., 2000; Maconochie et al., 1997; Manzanares et al., 1999a; Marshall et al., 1994; Nonchev et al., 1996b; Packer et al., 1998; Pöpperl et al., 1995; Studer et al., 1994; Zakany et al., 1997).

Cis-regulatory elements encompass enhancer, repressors, core promoters, matrix or scaffold attachment regions, and insulators (Arnone and Davidson, 1997; Pennacchio and Rubin, 2001). If these regulatory elements are conserved between two related (orthologous) genomic regions of different species, it is more likely that the conserved region is functional.

However, this methology has certain limitations. Finding regulatory elements is difficult, since the binding sites are short (8-10bp) and not always entirely conserved (Lenhard et al., 2003; Tümpel et al., 2002). The distance of the enhancer to the promoter is not essential for its regulatory properties; therefore regulatory regions can be located at considerable distances up- or downstream of the transcription start site. Also, there is no constraint on the location of the regulatory elements; it can even be positioned in the coding sequence (see Chapter 4). Specific genes can be regulated differently in various species, which means that certain regulatory elements of a regulatory module can be organized in a species-specific manner or even absent. The regulatory elements of duplicated genes may have diverged during evolution, such that each of the duplicated genes has only a subset of the regulatory elements of the original gene(McClintock et al., 2001).

A consensus binding sequence for a transcription factor has to be used cautiously for sequence comparisons, as a large number of such motifs can occur randomly in the genome and the vast majority of these have no role to play in gene regulation. Furthermore, sequence searches are only as good as the databases of transcription factor binding sites used to interrogate the DNA sequence itself. Many transcription factor-binding sites have yet to be described or are so degenerate that they are not useful unless proven by binding tests. Progress in defining and depositing new binding sites, and methods to test and calculate weighted averages in

the sites, will help in making sequence searches more productive. The development of new bioinformatic approaches that consider cooperatively between sites, probe non-linear relationships between motifs and compare relationships between coordinately regulated genes or paralogs to identify potential target sites is a major challenge and will be essential to exploit the emerging wealth of sequence information in non-coding regions (Fickett and Wasserman, 2000; GuhaThakurta and Stormo, 2001; Guigo et al., 2000; Hertz and Stormo, 1999; Stormo and Fields, 1998; Stormo, 2000; Weaver et al., 1999; Workman and Stormo, 2000).

9.2 Identification of regulatory regions using sequence alignments

Genomic comparison may lead to the identification of a highly conserved region containing potential *cis*-regulatory elements. By applying this method I was able to identify a highly conserved region in the *Hoxa2* intron (Chapter 3). I showed that this region contains the r4 enhancer by using chick electroporation and mouse transgenic analysis. The finding that the r4 enhancer is located in the *Hoxa2* intron was rather surprising, since the r4 enhancer of the paralogous member of *Hoxa2, Hoxb2*, is positioned in the intergenic region.

Even more surprising was the observation that the r2 enhancer of *Hoxa2* is located in the second exon (Chapter 4). In this case, applying comparative genomics was of limited use, since the r2 enhancer elements are located in the coding sequence. Phylogenic footprinting could not be used to identify this enhancer due to the high degree of conservation already imposed by the coding sequence. I had to perform deletion analysis to identify the r2 elements. Identifying the elements allowed me to perform local alignments of each region that has an important input in r2 regulation. This approach, made it simpler to determine the precise location of two ACAAT motifs, which are positioned adjacent to each other and exhibit an extremely high degree of conservation. However, the sequences of the RTE elements of different species show small variations, in particular, the RTE2 element. These changes do not appear to modulate species-specific expression, since r2 enhancers from different species were equally efficient in interspecies experiments. Further, aligning orthologous nucleotide and the

encoded protein sequences of the r2 elements, gave important hints about the way the r2 elements evolved in the coding sequence.

9.3 Genes can be regulated in a species-specific manner

In some species, *cis*-regulatory elements in a conserved region can be diverged or be absent, which could reflect species-specific regulation. The *Hoxa2* r4 module consists of three Hox/Pbx (PH) sites and one Prep/Meis (PM) site. The PM element shows an extremely high degree of conservation, with no variation among the analyzed species (Chapter 3). The second PH site is highly conserved among different species, although not identical. Intriguingly, the region between the PM and the PH site also exhibit a high degree of conservation. The first and third PH sites show a higher degree of divergence between amniotes and fishes. These two sites have an important input in r4 regulation in mice, whereas, in chick embryos, the input was important only to a limited degree. Overall, the conservation of these elements led to the identification of the r4 enhancer of *Hoxa2*, although some r4 elements are diverged. This could be due to species-specific differences in r4 regulation, which I observed in interspecies experiments.

The conserved region in the intron of *Hoxa2* was presence in all analyzed vertebrates, except in the frog orthologous genomic sequence, where I could not find any conservation, either in *Xenopus laevis* or in *tropicalis* (data not shown). This observation is intriguing, because *Hoxa2* is expressed in r4 in *Xenopus laevis* (Pasqualetti et al., 2000). Most likely, different regulatory mechanisms are responsible for r4 expression in frog, since I was unable to identify any PM or PH sites.

Relying only on comparison of orthologous genomic DNA of different mammals is not useful as the overall degree of conservation is too high to distinguish any functionally important regulatory regions. Hence, multiple species comparisons rather than pair-wise sets should be more fruitful in identifying conserved sequence strings or motifs that participate in regulation of gene expression. Including chick or fish genomic sequence in the alignment increases the chance that conserved sequences have a functional role. Therefore, in most cases, it is im-

portant to compare sequences with maximal phylogenetic distance. However, relying only on regions that are conserved between all vertebrates ignores the fact that species-specific control elements have evolved. I observed this phenomenon in regulatory studies on the Hoxb1 repressor region (Chapter 7). Here, I identified a repressor region in the intergenic region of *Hoxb1/b2* that mediates the downregulation of *Hoxb1* expression in the BA2. I started with deletion analysis, and, in sequential steps, I used comparative sequence analysis to locate the repressor region. This repressor region is present only in mammals, which is consistent with the observation that *Hoxb1* expression is not downregulated in the BA2 in chick.

In this case, identification of the conserved region would have been impossible without performing prior deletion experiments. I narrowed down the region to a smaller region (1.3kb) and searched within this small fragment for conserved regions. I was able to identify two conserved elements (CRI and CRII), the first of which exhibited repressor activity.

9.4 Pufferfish as a model system for genomic comparison

The pufferfish, *Fugu rubrides*, has been chosen as an attractive model system for genomic analysis, due to its compact genome. Its genome size is about one-eighth the size of the human genome (Brenner et al., 1993). This is not the case for the *Hox* cluster. In my analysis, the fugu *Hoxb1/b2* and *Hoxa2/a3* intergenic regions were not smaller compared to other species (see Chapters 3 and 6).

One must also be cautious about comparing the intergenic regions of duplicated genes, especially in ray fin fishes, which include the model systems medaka, zebrafish, and fugu (Amores et al., 1998; Amores et al., 2004; Naruse et al., 2000; Scemama et al., 2002). The function of the duplicated genes may have changed by neo- or sub-functionalization, which can have a profound effect on the conservation of each of the regulatory modules (Furutani-Seiki and Wittbrodt, 2004; McClintock et al., 2002; Prince and Pickett, 2002). In this case expression analysis of each of the duplicated gene is essential; the expression pattern can provide insights whether regulatory features were lost or added.

I observed this phenomenon in the *Hoxa2* rhombomeric regulatory regions discussed in Chapter 5. I used sequence alignments to evaluate the basis of the differential expression of the two co-paralogous genes *Hoxa2(a)* and *(b)* in fugu. Performing local comparisons among different species was extremely helpful in this regard. I showed that subtle changes in several elements are responsible for the differential expression of the two duplicated genes. Also, including the orthologous sequences of the duplicated *Hoxa2* genes of medaka helped to identify the sequence drift. Comparing the sequences and genomic organization of the *Hox* genes among different species was also helpful to determine at which phylogenetic stage, the changes occurred (Amores et al., 2004).

9.5 Searching for specific binding sites in conserved regions

Another approach of analyzing regulatory circuits would be to search for specific transcription factor binding sites, which have been implicated in regulating a specific target gene. Using this approach, I identified a highly conserved Krox20 binding site in the *Hoxb1/b2* intergenic region (Chapter 7). I showed that this region has an important role in restricting *Hoxb1* expression to r4 and is able to bind Krox20. In this project, I was actually searching for potential Krox20 binding sites in conserved regions, since Krox20 has been implicated in downregulating *Hoxb1* expression (Giudicelli et al., 2001). Therefore, I started searching not for conserved regions, but for potential Krox20 binding sites within conserved regions. The Krox20 site was located adjacent to a previously identified r3/5 repressor element (Studer et al., 1994); this strongly suggests that the Krox20 binding site was real.

9.6 Diversity in the *Hoxa2* elements as a way of generating changes in morphology

In my thesis, I used sequence conservation as a landmark to identify regulatory regions. However, knowing the precise location and composition of different *cis*-regulatory regions of *trans*-acting regulators can be useful to detect variations within the regulatory sequence which could be linked to altered regulatory output. It has been argued that morpho-

logical diversity during evolution are due to changes in sequences in *cis*-regulatory regions (Belting et al., 1998; Carroll, 2000; Weatherbee et al., 1998).

Hoxa2 plays an important role during craniofacial development, by acting as a selector gene, directing the morphological fate of second BA neural crest cells (Gendron-Maguire et al., 1993; Pasqualetti et al., 2000; Rijli et al., 1993). It may be possible to correlate sequence drift present in known *Hoxa2* NCC elements to craniofacial diversity during vertebrate evolution (Maconochie et al., 1999; Tümpel et al., 2002).

I started to address this question by performing sequence comparisons of the *Hoxa2* regulatory region in different dog breeds (subspecies). I have access to genomic DNA of 85 different dog breeds and, since they all show different craniofacial-morphological diversity, it may be possible to make correlations between the sequence of *Hoxa2* control regions and specific craniofacial structures (data not shown). I subdivided the *Hoxa2* regulatory region into seven overlapping regions and I cloned each region by PCR to systematically generate the entire contig from each breed. This region contains all known *Hoxa2* enhancers, including the r3/r5, NCC, r4, and r2 enhancers. They were sequenced and analyzed by bioinformatic approaches. Preliminary data suggest that there are subtle changes in the r3/5 enhancer region. Functional comparisons have to be performed to draw correlations between sequence drift and morphological changes. However, changes identified in the regulatory regions do not have to be the primary cause of morphological changes, but can be genetically linked to another uncharacterized mutation or sequence drift. Therefore, this analysis can be very difficult and time-consuming.

Another approach was to analyze the bat *Hoxa2* regulatory regions. Bats show a variety of craniofacial structures. I investigated whether the bat *Hoxa2* NCC enhancer elements direct different reporter expression in transgenic mice. I have isolated the genomic locus of *Hoxa2* of the short-tailed fruit bat by phage screening, cloned the r3/5 and NCC enhancer, linked it to a *lacZ* reporter construct and performed pronuclear injection in mice. The reporter expression looked very similar to that seen with the murine control region (data not shown and

see Nonchev et al., 1996b). Further comparative analysis has to be performed to precisely examine whether these two patterns are identical, or whether they differ to some degree.

9.7 Identifying *trans*-acting factors controlling expression of *Hoxa2* and *Hoxb1*

In my thesis, I identified several *cis*-regulatory elements which control expression of *Hoxa2* and *Hoxb1*. I was able to show that r4 expression of *Hoxa2* is mediated by Hoxb1 and its co-factors Prep, Meis and Pbx (Chapter 3). I present data, which shows that restriction of *Hoxb1* expression in r3/5 is mediated by the zinc-finger transcription factor Krox20 (Chapter 6).

However, in a couple of cases, I was unable to identify the *trans*-acting factors binding to the newly identified elements; the *Hoxa2* r2 enhancer consists of several elements, which includes two highly conserved ACAAT motifs. Although the expression pattern of many transcription factors has been analyzed in the hindbrain, it still needs to be determined which transcription factors regulate the expression of *Hoxa2* in r2.

There are several ways to approach this question. One would be using *Drosophila*, in which homeotic genes have been originally identified (Lewis, 1978). *Drosophila* is an excellent genetic system for understanding regulatory genetic networks (Levine and Davidson, 2005; Markstein et al., 2002). *Hox* genes show a remarkable phylogenetic conservation regarding their sequence and also their physical location along the chromosomes (Krumlauf, 1994). Therefore, *Drosophila* as a model system can be useful in understanding the regulation of genes of vertebrates (Gould et al., 1997; Lutz et al., 1996; Maconochie et al., 1997; Pöpperl et al., 1995).

The *Drosophila* homolog of *Hoxa2*, *proboscipedia*, is expressed in gnathal segments. It is one of the most anteriorly expressed *Hox* genes in *Drosophila.* Proboscipedia (Pb) has been shown to be required for tarsus, maxillary palp, and proboscis determination (Percival-Smith et al., 1997). It has been tested whether the mouse Hoxa2 and Pb are functional equivalents in *Drosophila* (Percival-Smith and Bondy, 1999). Ectopic expression of *Hoxa2* was able

to inhibit Sex Combs Reduced activity, but a complete rescue of pb phenotypes were not observed (Percival-Smith and Bondy, 1999). This suggests that Hoxa2 shares only partial activity of Pb. The r2 enhancer of *Hoxa2* has been also tested in *Drosophila* (Frasch et al., 1995). Transgenic flies were generated with the murine *Hoxa2* r2 enhancer linked to a reporter gene. Reporter expression was visible in the maxillary and labial segments of the embryonic head. Pb also shows expression in these domains, however, *Pb* and *Hoxa2* r2 directed reporter expression in *Drosophila* also shows much variation, especially along the dorsoventral axis (Frasch et al., 1995). A 1kb fragment encompassing enhancer elements directing reporter expression, which is very similar to the endogenous gene expression of *Pb*, has been identified in the second intron of the proboscipedia locus (Randazzo et al., 1991). This suggests that the location of the *cis*-regulatory has changed during evolution, but that the same *trans*-acting factors are involved in regulating *Pb* expression and the anterior expression of *Hoxa2*. The intron regions contains one ACAAT motif which is a highly conserved between *Drosophila melanogaster* and *Drosophila pseudoobscura* (data not shown). It is possible that related upstream factors are responsible for the similar expression pattern of these two orthologous genes and that regulatory elements are conserved. It would be possible to test whether the factors shown to regulate *Pb* expression also are involved in regulating *Hoxa2* expression in r2. It would be possible to establish a stable transgenic line of flies carrying the *Hoxa2* r2/*lacZ* construct and cross it with various mutant flies.

Another way would be to apply biochemical approaches, e.g. yeast one-hydrid screening would be a feasible way to identify the transcription factors interacting with the ACAAT or RTE elements of the r2 enhancer of *Hoxa2*. Another option would be to fractionate nuclear extract and test whether oligonucleotides spanning any of the r2 elements show specific binding in EMSA experiments. After several purification steps it might be possible to obtain sufficient pure fraction containing specific binding factors for mass spectrometric analysis.

It is also unknown which *trans*-acting factors mediate the repression of *Hoxb1* expression in BA2 (Chapter 7). The region of conservation is relatively large (~150bp), which makes

it difficult to find any obvious single transcription factor binding sites. For this purpose the repressor region has to be further analyzed by deletion experiments. In this case focusing on mouse as a modelsystem would be convenient. If the region is narrowed down one could search for potential binding site, which could be confirmed by biochemical methods as discussed above. Another approach would be to search for candidate genes, which are expressed at high levels in BA2 and which have been shown to exhibit repressor activity or which are known to interact with co-repressors.

9.8 Summary

Overall, comparative sequence analysis is a powerful tool for identifying *cis*-regulatory elements and subtle changes in regulatory regions. However, this technique is not always applicable as discussed above. The only alternative to identify *cis*-regulatory regions would have been to perform extensive deletion analysis. Not only is the comparative approach faster and less costly, but using sequence alignments provides additional information about the divergence and conservation of *cis*-regulatory elements during evolution.

References

Akam, M. (1987) The molecular basis for metameric pattern in the Drosophila embryo. *Development* **101,** 1-22.

Albano, R.M., Arkell, R., Beddington, R.S., and Smith, J.C. (1994) Expression of inhibin subunits and follistatin during postimplantation mouse development: decidual expression of activin and expression of follistatin in primitive streak, somites and hindbrain. *Development* **120,** 803-13.

Amores, A., Force, A., Yan, Y.-L., Joly, L., Amemiya, C., Fritz, A., Ho, R., Langeland, J., Prince, V., Wang, Y.-L., Westerfield, M., Ekker, M., and Postlehwait, J. (1998) Zebrafish *hox* clusters and vertebrate genome evolution. *Science* **282,** 1711-1714.

Amores, A., Suzuki, T., Yan, Y.L., Pomeroy, J., Singer, A., Amemiya, C., and Postlethwait, J.H. (2004) Developmental roles of pufferfish Hox clusters and genome evolution in ray-fin fish. *Genome Res.* **14,** 1-10.

Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P.W.J., Krumlauf, R., and Brenner, S. (1995) Detecting conserved regulatory elements with the model genome of the Japanese puffer fish *Fugu rubripes*. *Proc. Natl. Acad. Sci. USA* **92,** 1684-1688.

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.M., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., Gelpke, M.D., Roach, J., Oh, T., Ho, I.Y., Wong, M., Detter, C., Verhoef, F., Predki, P., Tay, A., Lucas, S., Richardson, P., Smith, S.F., Clark, M.S., Edwards, Y.J., Doggett, N., Zharkikh, A., Tavtigian, S.V., Pruss, D., Barnstead, M., Evans, C., Baden, H., Powell, J., Glusman, G., Rowen, L., Hood, L., Tan, Y.H., Elgar, G., Hawkins, T., Venkatesh, B., Rokhsar, D., and Brenner, S. (2002) Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. *Science* **297,** 1301-10.

Arenkiel, B.R., Tvrdik, P., Gaufo, G.O., and Capecchi, M.R. (2004) Hoxb1 functions in both motoneurons and in tissues of the periphery to establish and maintain the proper neuronal circuitry. *Genes Dev.* **18,** 1539-52.

Arnone, M.I., and Davidson, E.H. (1997) The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124,** 1851-1864.

Bailey, W.J., Kim, J., Wagner, G.P., and Ruddle, F.H. (1997) Phylogenetic reconstruction of vertebrate Hox cluster duplications. *Mol Biol Evol* **14,** 843-53.

Balling, R., Mutter, G., Gruss, P., and Kessel, M. (1989) Craniofacial abnormalities induced by ectopic expression of the homeobox gene *Hox-1.1* in transgenic mice. *Cell* **58,** 337-347.

Barrow, J., and Capecchi, M. (1996) Targeted disruption of the *Hoxb2* locus in mice interferes with expression of *Hoxb1* and *Hoxb4*. *Development* **122,** 3817-3828.

Barrow, J.R., Stadler, H.S., and Capecchi, M.R. (2000) Roles of Hoxa1 and Hoxa2 in patterning the early hindbrain of the mouse. *Development* **127,** 933-44.

Begbie, J., Brunet, J.F., Rubenstein, J.L., and Graham, A. (1999) Induction of the epibranchial placodes. *Development* **126,** 895-902.

Bell, E., Wingate, R.J., and Lumsden, A. (1999) Homeotic transformation of rhombomere identity after localized *Hoxb1* misexpression. *Science* **284,** 21682171.

Belting, H.G., Shashikant, C.S., and Ruddle, F.H. (1998) Modification of expression and cis-regulation of Hoxc8 in the evolution of diverged axial morphology. *Proc Natl Acad Sci U S A* **95,** 2355-60.

Bel-Vialar, S., Itasaki, N., and Krumlauf, R. (2002) Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. *Development* **129,** 5103-15.

Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F., and Blasi, F. (1998) The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. *EMBO J.* **17,** 1434-45.

Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., Haussler, D., and Miller, W. (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14,** 708-15.

Bolos, V., Peinado, H., Perez-Moreno, M.A., Fraga, M.F., Esteller, M., and Cano, A. (2003) The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci* **116,** 499-511.

Bonfield, J.K., Smith, K., and Staden, R. (1995) A new DNA sequence assembly program. *Nuc. Acids Res.* **23,** 4992-4999.

Brend, T., Gilthorpe, J., Summerbell, D., and Rigby, P.W. (2003) Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of Hoxb4 expression. *Development* **130,** 2717-28.

Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B., and Aparicio, S. (1993) Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* **366,** 265-268.

Bronner-Fraser, M. (2004) Development. Making sense of the sensory lineage. *Science* **303,** 966-8.

Burglin, T.R. (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res* **25,** 4173-80.

Cambronero, F., and Puelles, L. (2000) Rostrocaudal nuclear relationships in the avian medulla oblongata: a fate map with quail chick chimeras. *J Comp Neurol* **427,** 522-45.

Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000) The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* **2,** 76-83.

Capecchi, M.R. (1989) The new mouse genetics: altering the genome by gene targeting. *Trends Genet* **5,** 70-76.

Carroll, S.B. (1995) Homeotic genes and the evolution of arthropods and chordates. *Nature* **376,** 479-485.

Carroll, S.B. (2000) Endless forms: the evolution of gene regulation and morphological diversity. *Cell* **101,** 577-80.

Chambeyron, S., and Bickmore, W.A. (2004) Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* **18,** 1119-30.

Chan, S.K., Ryoo, H.D., Gould, A., Krumlauf, R., and Mann, R.S. (1997) Switching the *in vivo* specificity of a minimal *HOX*-responsive element. *Development* **124,** 2007-2014.

Chan, S.-K., and Mann, R.S. (1996) A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc. Natl. Acad. Sci. U S A* **93,** 5223-5228.

Chang, C.P., Brocchieri, L., Shen, W.F., Largman, C., and Cleary, M.L. (1996) Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the *Hox* locus. *Mol. Cell. Biol.* **16,** 1734-1745.

Chang, C.P., Jacobs, Y., Nakamura, T., Jenkins, N.A., Copeland, N.G., and Cleary, M.L. (1997) Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol. Cell. Biol.* **17,** 5679-87.

Chavrier, P., Vesque, C., Galliot, B., Vigneron, M., Dollé, P., Duboule, D., and Charnay, P. (1990) The segment-specific gene *Krox-20* encodes a transcription factor with binding sites in the promoter of the *Hox 1.4* gene. *EMBO J.* **9,** 1209-1218.

Chen, J., and Ruley, H. (1998) An enhancer element in the *EphA2 (Eck)* gene sufficient for rhombomere-specific expression is activated by *Hoxa1* and *Hoxb1* homeobox proteins. *Journal of Biological Chemistry* **273,** 24670-24675.

Chisaka, O., and Capecchi, M.R. (1991) Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene *hox1.5*. *Nature* **350,** 473-479.

Chisaka, O., Musci, T.S., and Capecchi, M.R. (1992) Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene *Hox-1.6*. *Nature* **355,** 516-520.

Cho, K., Morita, E., Wright, C., and De Robertis, E. (1991) Over-expression of a homeodomain protein confers axis-forming activity to uncommitted *Xenopus* embryonic cells. *Cell* **65,** 55-64.

Condie, B.G., and Capecchi, M.R. (1993) Mice homozygous for a targeted disruption of *Hoxd-3(Hox-4.1)* exhibit anterior transformations of the first and second cervical vertebrae, the atlas and axis. *Development* **119,** 579-595.

Condie, B.G., and Capecchi, M.R. (1994) Mice with targeted disruptions in the paralogous genes *Hoxa-3* and *Hoxd-3* reveal synergistic interactions. *Nature* **370,** 304-307.

Cordes, S.P., and Barsh, G.S. (1994) The mouse segmentation gene *kr* encodes a novel basic domain-leucine zipper transcription factor. *Cell* **79,** 1025-1034.

Couly, G., and Le Douarin, N.M. (1990) Head morphogenesis in embryonic avian chimeras: evidence for a segmental pattern in the ectoderm corresponding to the neuromeres. *Development* **108,** 543-558.

Couly, G., Creuzet, S., Bennaceur, S., Vincent, C., and Le Douarin, N.M. (2002) Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development* **129,** 1061-73.

Davenne, M., Maconochie, M.K., Neun, R., Pattyn, A., Chambon, P., Krumlauf, R., and Rijli, F.M. (1999) *Hoxa2* and *Hoxb2* control dorsoventral patterns of neuronal development in the rostral hindbrain. *Neuron* **22,** 677-691.

de The, H., Vivanco-Ruiz, M.M., Tiollais, P., Stunnenberg, H., and Dejean, A. (1990) Identification of a retinoic acid responsive element in the retinoic acid receptor beta gene. *Nature* **343,** 177-80.

Dekker, E.J., Pannese, M., Houtzager, E., Boncinelli, E., and Durston, A. (1992) Colinearity in the *Xenopus laevis Hox-2* complex. *Mech Dev* **40(1),** 3-12.

Deol, M.S. (1964) The abnormalities of the inner ear in *kreisler* mice. *J Embryol Exp Morphol* **12,** 475-490.

Di Rocco, G., Gavalas, A., Popperl, H., Krumlauf, R., Mavilio, F., and Zappavigna, V. (2001) The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. *J. Biol. Chem.* **1,** 1.

Dollé, P., Izpisua-Belmonte, J.C., Falkenstein, H., Renucci, A., and Duboule, D. (1989a) Coordinate expression of the murine Hox-5 complex homoeobox-containing genes during limb pattern formation. *Nature* **342,** 767-72.

Dollé, P., Izpisùa-Belmonte, J.C., Falkenstein, H., Renucci, A., and Duboule, D. (1989b) Co-ordinate expression of the murine *Hox-5* complex homeobox-containing genes during limb pattern formation. *Nature* **342,** 767-772.

Dollé, P., Ruberte, E., Leroy, P., Morriss-Kay, G., and Chambon, P. (1990) Retinoic acid receptors and cellular retinoid binding proteins. I. A systematic study of their differential pattern of transcription during mouse organogenesis. *Development* **110,** 1133-1151.

Dressler, G.R., and Gruss, P. (1989) Anterior boundaries of *Hox* gene expression in mesoderm-derived structures correlate with the linear gene order along the chromosome. *Differentiation* **41,** 193-201.

Duboule, D., and Morata, G. (1994) Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* **10,** 358-364.

Duester, G. (2000) Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid. *Eur J Biochem* **267,** 4315-24.

Dupe, V., and Lumsden, A. (2001) Hindbrain patterning involves graded responses to retinoic acid signalling. *Development* **128,** 2199-208.

Dupé, V., Davenne, M., Brocard, J., Dollé, P., Mark, M., Dierich, A., Chambon, P., and Rijli, F. (1997) *In vivo* functional analysis of the *Hoxa1* 3' retinoid response element (3' RARE). *Development* **124,** 399-410.

Dupé, V., Ghyselinck, N.B., Wendling, O., Chambon, P., and Mark, M. (1999) Key roles of retinoic acid receptors alpha and beta inthe patterning of the caudal hindbrain, pharyngeal arches and otocyst in the mouse. *Development* **126,** 5051-5059.

Ellies, D.L., Church, V., Francis-West, P., and Lumsden, A. (2000) The WNT antagonist cSFRP2 modulates programmed cell death in the developing hindbrain. *Development* **127,** 5285-95.

Ellies, D.L., Tucker, A.S., and Lumsden, A. (2002) Apoptosis of premigratory neural crest cells in rhombomeres 3 and 5: consequences for patterning of the branchial region. *Dev. Biol.* **251,** 118-28.

Farlie, P.G., Kerr, R., Thomas, P., Symes, T., Minichiello, J., Hearn, C.J., and Newgreen, D. (1999) A paraxial exclusion zone creates patterned cranial neural crest cell outgrowth adjacent to rhombomeres 3 and 5. *Dev. Biol.* **213,** 70-84.

Ferretti, E., Schulz, H., Talarico, D., Blasi, F., and Berthelsen, J. (1999) The PBX-regulating protein PREP1 is present in different PBX-complexed forms in mouse. *Mech Dev* **83,** 53-64.

Ferretti, E., Marshall, H., Pöpperl, H., Maconochie, M., Krumlauf, R., and Blasi, F. (2000) Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. *Development* **127,** 155-166.

Ferrier, D.E., Minguillon, C., Holland, P.W., and Garcia-Fernandez, J. (2000) The amphioxus Hox cluster: deuterostome posterior flexibility and Hox14. *Evol Dev* **2,** 284-93.

Fickett, J.W., and Wasserman, W.W. (2000) Discovery and modeling of transcriptional regulatory regions. *Curr Opin Biotechnol* **11,** 19-24.

Finnerty, J.R., Pang, K., Burton, P., Paulson, D., and Martindale, M.Q. (2004) Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. *Science* **304,** 1335-7.

Fox, E.A. (2000) The previously identified r3/r5 repressor may require the cooperation of additional negative elements for rhombomere restriction of Hoxb1. *Brain Res Dev Brain Res* **120,** 151-64.

Frasch, M., Chen, X., and Lufkin, T. (1995) Evolutionary-conserved enhancers direct region-specific expression of the murine *Hoxa-1* and *Hoxa-2* loci in both mice and *Drosophila*. *Development* **121,** 957-974.

Fraser, S., Keynes, R., and Lumsden, A. (1990) Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* **344,** 431-435.

Frohman, M., and Martin, G. (1992) Isolation and analysis of embryonic expression of Hox-4.9, a member of the murine labial-like gene family. *Mech. Dev.* **38,** 55-67.

Frohman, M.A., Boyle, M., and Martin, G.R. (1990) Isolation of the mouse *Hox-2.9* gene; analysis of embryonic expression suggests that positional information along the anterior-posterior axis is specified by mesoderm. *Development* **110,** 589-607.

Frohman, M.A., Martin, G.R., Cordes, S.P., Halamek, L.P., and Barsh, G.S. (1993) Altered rhombomere-specific gene expression and hyoid bone differentiation in the mouse segmentation mutant, *kreisler (kr)*. *Development* **117,** 925-936.

Furutani-Seiki, M., and Wittbrodt, J. (2004) Medaka and zebrafish, an evolutionary twin study. *Mech Dev* **121,** 629-37.

Gale, E., Zile, M., and Maden, M. (1999) Hindbrain respecification in the retinoid-deficient quail. *Mech Dev* **89,** 43-54.

Garcia-Castro, M.I., Marcelle, C., and Bronner-Fraser, M. (2002) Ectodermal Wnt function as a neural crest inducer. *Science* **13,** 13.

Gaufo, G.O., Thomas, K.R., and Capecchi, M.R. (2003) Hox3 genes coordinate mechanisms of genetic suppression and activation in the generation of branchial and somatic motoneurons. *Development* **130,** 5191-201.

Gavalas, A., Davenne, M., Lumsden, A., Chambon, P., and Rijli, F.M. (1997) Role of *Hoxa-2* in axon pathfinding and rostral hindbrain patterning. *Development* **124,** 3693-3702.

Gavalas, A., Studer, M., Lumsden, A., Rijli, F.M., Krumlauf, R., and Chambon, P. (1998) *Hoxa1* and *Hoxb1* synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch. *Development* **125,** 1123-1136.

Gavalas, A., Trainor, P., Ariza-McNaughton, L., and Krumlauf, R. (2001) Synergy between Hoxa1 and Hoxb1: the relationship between arch patterning and the generation of cranial neural crest. *Development* **128,** 3017-27.

Gavalas, A. (2002) ArRAnging the hindbrain. *Trends Neurosci* **25,** 61-4.

Gavalas, A., Ruhrberg, C., Livet, J., Henderson, C.E., and Krumlauf, R. (2003) Neuronal defects in the hindbrain of Hoxa1, Hoxb1 and Hoxb2 mutants reflect regulatory interactions among these Hox genes. *Development.*

Geada, A.M.C., Gaunt, S.J., Azzawi, M., Shimeld, S.M., Pearce, J., and Sharpe, P.T. (1992) Sequence and embryonic expression of the murine Hox-3.5 gene. *Development* **116,** 497-506.

Gehring, W., Muller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y., Otting, G., and Wuthrich, K. (1990) The structure of the homeodomain and its functional implications. *TIG* **6,** 323-329.

Gendron-Maguire, M., Mallo, M., Zhang, M., and Gridley, T. (1993) Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell* **75,** 1317-1331.

Ghislain, J., Desmarquet-Trin-Dinh, C., Gilardi-Hebenstreit, P., Charnay, P., and Frain, M. (2003) Neural crest patterning: autoregulatory and crest-specific elements cooperate for Krox20 transcriptional control. *Development* **130,** 941-53.

Gibbs, R.A., Weinstock, G.M., Metzker, M.L., Muzny, D.M., Sodergren, E.J., Scherer, S., Scott, G., Steffen, D., Worley, K.C., Burch, P.E., Okwuonu, G., Hines, S., Lewis, L., DeRamo, C., Delgado, O., Dugan-Rocha, S., Miner, G., Morgan, M., Hawes, A., Gill, R., Celera, Holt, R.A., Adams, M.D., Amanatides, P.G., Baden-Tillson, H., Barnstead, M., Chin, S., Evans, C.A., Ferriera, S., Fosler, C., Glodek, A., Gu, Z., Jennings, D., Kraft, C.L., Nguyen, T., Pfannkoch, C.M., Sitter, C., Sutton, G.G., Venter, J.C., Woodage, T., Smith, D., Lee, H.M., Gustafson, E., Cahill, P., Kana, A., Doucette-Stamm, L., Weinstock, K., Fechtel, K., Weiss, R.B., Dunn, D.M., Green, E.D., Blakesley, R.W., Bouffard, G.G., De Jong, P.J., Osoegawa, K., Zhu, B., Marra, M., Schein, J., Bosdet, I., Fjell, C., Jones, S., Krzywinski, M., Mathewson, C., Siddiqui, A., Wye, N., McPherson, J., Zhao, S., Fraser, C.M., Shetty, J., Shatsman, S., Geer, K., Chen, Y., Abramzon, S., Nierman, W.C., Havlak, P.H., Chen, R., Durbin, K.J., Egan, A., Ren, Y., Song, X.Z., Li, B., Liu, Y., Qin, X., Cawley, S., Cooney, A.J., D'Souza, L.M., Martin, K., Wu, J.Q., Gonzalez-Garay, M.L., Jackson, A.R., Kalafus, K.J., McLeod, M.P., Milosavljevic, A., Virk, D., Volkov, A., Wheeler, D.A., Zhang, Z., Bailey, J.A., Eichler, E.E., Tuzun, E., et al. (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428,** 493-521.

Giudicelli, F., Taillebourg, E., Charnay, P., and Gilardi-Hebenstreit, P. (2001) Krox-20 patterns the hindbrain through both cell-autonomous and non cell-autonomous mechanisms. *Genes Dev.* **15,** 567-80.

Giudicelli, F., Gilardi-Hebenstreit, P., Mechta-Grigoriou, F., Poquet, C., and Charnay, P. (2003) Novel activities of Mafb underlie its dual role in hindbrain segmentation and regional specification. *Dev. Biol.* **253,** 150-62.

Glass, C.K., and Rosenfeld, M.G. (2000) The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* **14,** 121-41.

Goddard, J., Rossel, M., Manley, N., and Capecchi, M. (1996) Mice with targeted disruption of *Hoxb1* fail to form the motor nucleus of the VIIth nerve. *Development* **122,** 3217-3228.

Godsave, S.F., Koster, C.H., Getahun, A., Mathu, M., Hooiveld, M., van der Wees, J., Hendriks, J., and Durston, A.J. (1998) Graded retinoid responses inthe developing hindbrain. *Developmental Dynamics* **213,** 39-49.

Golding, J.P., Trainor, P., Krumlauf, R., and Gassmann, M. (2000) Defects in pathfinding by cranial neural crest cells in mice lacking the Neuregulin receptor ErbB4. *Nature Cell Biology* **2,** 103-109.

Gould, A. (1997) Functions of mammalian *Polycomb*-group and *trithorax*-group related genes. *Curr. Opin. Genet. Dev.* **7,** 488-494.

Gould, A., Morrison, A., Sproat, G., White, R.A., and Krumlauf, R. (1997) Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping *Hox* expression patterns. *Genes and Development* **11,** 900-913.

Gould, A., Itasaki, N., and Krumlauf, R. (1998) Initiation of rhombomeric *Hoxb4* expression requires induction by somites and a retinoid pathway. *Neuron* **21,** 39-51.

Graham, A., Papalopulu, N., and Krumlauf, R. (1989) The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* **57,** 367-378.

Graham, A., Heyman, I., and Lumsden, A. (1993) Even-numbered rhombomeres control the apoptotic elimination of neural crest cells from odd-numbered rhombomeres in the chick hindbrain. *Development* **119,** 233-245.

Graham, A., Francis-West, P., Brickell, P., and Lumsden, A. (1994) The signalling molecule BMP4 mediates apoptosis in the rhombencephalic neural crest. *Nature* **372,** 684-686.

Grammatopoulos, G.A., Bell, E., Toole, L., Lumsden, A., and Tucker, A.S. (2000) Homeotic transformation of branchial arch identity after Hoxa2 overexpression. *Development* **127,** 5355-65.

GuhaThakurta, D., and Stormo, G.D. (2001) Identifying target sites for cooperatively binding factors. *Bioinformatics* **17,** 608-21.

Guigo, R., Agarwal, P., Abril, J.F., Burset, M., and Fickett, J.W. (2000) An assessment of gene prediction accuracy in large DNA sequences. *Genome Res.* **10,** 1631-42.

Guthrie, S., and Lumsden, A. (1991) Formation and regeneration of rhombomere boundaries in the developing chick hindbrain. *Development* **112,** 221-229.

Haerry, T., and Gehring, W. (1996) Intron of the mouse *Hoxa-7* gene contains conserved homeodomain binding sites that can function as an enhancer element in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **93,** 13884-13889.

Helmbacher, F., Pujades, C., Desmarquet, C., Frain, M., Rijli, F.M., Chambon, P., and Charnay, P. (1998) *Hoxa1* and *Krox20* synergize to control the development of rhombomere 3. *Development* **125,** 4739-4748.

Hertz, G.Z., and Stormo, G.D. (1999) Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15,** 563-77.

Heyman, R.A., Mangelsdorf, D.J., Dyck, J.A., Stein, R.B., Eichele, G., Evans, R.M., and Thaller, C. (1992) 9-cis retinoic acid is a high affinity ligand for the retinoid X receptor. *Cell* **68,** 397-406.

Hillier, L.W., Miller, W., Birney, E., Warren, W., Hardison, R.C., Ponting, C.P., Bork, P., Burt, D.W., Groenen, M.A., Delany, M.E., Dodgson, J.B., Chinwalla, A.T., Cliften, P.F., Clifton, S.W., Delehaunty, K.D., Fronick, C., Fulton, R.S., Graves, T.A., Kremitzki, C., Layman, D., Magrini, V., McPherson, J.D., Miner, T.L., Minx, P., Nash, W.E., Nhan, M.N., Nelson, J.O., Oddy, L.G., Pohl, C.S., Randall-Maher, J., Smith, S.M., Wallis, J.W., Yang, S.P., Romanov, M.N., Rondelli, C.M., Paton, B., Smith, J., Morrice, D., Daniels, L., Tempest, H.G., Robertson, L., Masabanda, J.S., Griffin, D.K., Vignal, A., Fillon, V., Jacobbson, L., Kerje, S., Andersson, L., Crooijmans, R.P., Aerts, J., van der Poel, J.J., Ellegren, H., Caldwell, R.B., Hubbard, S.J., Grafham, D.V., Kierzek, A.M., McLaren, S.R., Overton, I.M., Arakawa, H., Beattie, K.J., Bezzubov, Y., Boardman, P.E., Bonfield, J.K., Croning, M.D., Davies, R.M., Francis, M.D., Humphray, S.J., Scott, C.E., Taylor, R.G., Tickle, C., Brown, W.R., Rogers, J., Buerstedde, J.M., Wilson, S.A., Stubbs, L., Ovcharenko, I., Gordon, L., Lucas, S., Miller, M.M., Inoko, H., Shiina, T., Kaufman, J., Salomonsen, J., Skjoedt, K., Wong, G.K., Wang, J., Liu, B., Yu, J., Yang, H., Nefedov, M., Koriabine, M., Dejong, P.J., Goodstadt, L., Webber, C., Dickens, N.J., Letunic, I., Suyama, M., Torrents, D., von Mering, C., Zdobnov, E.M., et al. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432,** 695-716.

Hoffmann, B., Lehmann, J.M., Zhang, X.K., Hermann, T., Husmann, M., Graupner, G., and Pfahl, M. (1990) A retinoic acid receptor-specific element controls the retinoic acid receptor-beta promoter. *Mol Endocrinol* **4,** 1727-36.

Horan, G.S., Wu, K., Wolgemuth, D.J., and Behringer, R.R. (1994) Homeotic transformation of cervical vertebrae in *Hoxa-4* mutant mice. *Proc. Natl. Acad. Sci. U S A* **91,** 12644-12648.

Horan, G.S., Ramirez-Solis, R., Featherstone, M.S., Wolgemuth, D.J., Bradley, A., and Behringer, R.R. (1995) Compound mutants for the paralogous *Hoxa-4, Hoxb-4*, and *Hoxd-4* genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. *Genes Dev.* **9,** 1667-1677.

Huang, D., Chen, S.W., Langston, A.W., and Gudas, L.J. (1998) A conserved retinoic acid responsive element in the murine *Hoxb-1* gene is required for expression in the developing gut. *Development* **125,** 3235-3246.

Hunt, P., Gulisano, M., Cook, M., Sham, M., Faiella, A., Wilkinson, D., Boncinelli, E., and Krumlauf, R. (1991) A distinct *Hox* code for the branchial region of the head. *Nature* **353,** 861-864.

Hunter, M.P., and Prince, V.E. (2002) Zebrafish hox paralogue group 2 genes function redundantly as selector genes to pattern the second pharyngeal arch. *Dev. Biol.* **247,** 367-89.

Inoue, T., and Krumlauf, R. (2001) An impulse to the brain: Using in vivo electroporation. *Nature Neuroscience* **4,** 6-8.

Irving, C., Nieto, M., Das Gupta, R., Charnay, P., and Wilkinson, D. (1996) Progressive spatial restriction of Sek1 and Krox20 gene expression during hindbrain segmentation. *Dev. Biol.* **173,** 26-38.

Irving, C., and Mason, I. (2000) Signalling by FGF8 from the isthmus patterns anterior hindbrain and establishes the anterior limit of *Hox* gene expression. *Development* **127,** 177-186.

Itasaki, N., Bel-Vialar, S., and Krumlauf, R. (1999) "Shocking" developments in chick embryology: electroporation and *in ovo* gene expression. *Nature Cell Biology* **1,** E203-207.

Izpisua-Belmonte, J., Falkenstein, H., Dollé, P., Renucci, A., and Duboule, D. (1991a) Murine genes related to the *Drosophila AbdB* homeotic gene are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10,** 2279-2289.

Izpisua-Belmonte, J.C., Falkenstein, H., Dollé, P., Renucci, A., and Duboule, D. (1991b) Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10,** 2279-89.

Jacobs, Y., Schnabel, C.A., and Cleary, M.L. (1999) Trimeric association of Hox and TALE homeodomain proteins mediates *Hoxb2* hindbrain enhancer activity. *Mol. Cell. Biol.* **19,** 5134-5142.

Jegalian, B.G., and De Robertis, E.M. (1992) Homeotic transformations in the mouse induced by overexpression of a human Hox3.3 transgene. *Cell* **71,** 901-910.

Kanzler, B., Kuschert, S.J., Liu, Y.-H., and Mallo, M. (1998) *Hoxa2* restricts the chondrogenic domain and inhibits bone formation during development of the branchial area. *Development* **125,** 2587-2597.

Kappen, C., Schughart, K., and Ruddle, F.H. (1989) Two steps in the evolution of antennapedia-class vertebrate homeobox genes. *Proc. Natl. Acad. Sci. U S A* **86,** 5459-5463.

Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J., Weber, R.J., Haussler, D., and Kent, W.J. (2003) The UCSC Genome Browser Database. *Nucleic Acids Res* **31,** 51-4.

Kastner, P., Mark, M., Ghyselinck, N., Krezel, W., Dupé, V., Grondona, J.M., and Chambon, P. (1997a) Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development* **124,** 313-326.

Kastner, P., Messaddeq, N., Mark, M., Wendling, O., Grondona, J.M., Ward, S., Ghyselinck, N., and Chambon, P. (1997b) Vitamin A deficiency and mutations of RXRalpha, RXRbeta and RARalpha lead to early differentiation of embryonic ventricular cardiomyocytes. *Development* **124,** 4749-58.

Kent, W.J. (2002) BLAT--the BLAST-like alignment tool. *Genome Res.* **12,** 656-64.

Kessel, M., and Gruss, P. (1990) Murine developmental control genes. *Science* **249,** 374-379.

Kessel, M., and Gruss, P. (1991) Homeotic transformations of murine prevertebrae and concommitant alteration of Hox codes induced by retinoic acid. *Cell* **67,** 89-104.

Kim, C.B., Amemiya, C., Bailey, W., Kawasaki, K., Mezey, J., Miller, W., Minoshima, S., Shimizu, N., Wagner, G., and Ruddle, F. (2000) Hox cluster genomics in the horn shark, Heterodontus francisci. *Proc Natl Acad Sci U S A* **97,** 1655-60.

Kliewer, S.A., Umesono, K., Mangelsdorf, D.J., and Evans, R.M. (1992) Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature* **355,** 446-449.

Kmita, M., Fraudeau, N., Herault, Y., and Duboule, D. (2002) Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* **420,** 145-50.

Kmita, M., and Duboule, D. (2003) Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301,** 331-3.

Köntges, G., and Lumsden, A. (1996) Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny. *Development* **122,** 3229-3242.

Krezel, W., Dupe, V., Mark, M., Dierich, A., Kastner, P., and Chambon, P. (1996) RXR gamma null mice are apparently normal and compound RXR alpha +/-/RXR beta -/-/RXR gamma -/- mutant mice are viable. *Proc Natl Acad Sci U S A* **93,** 9010-4.

Krumlauf, R. (1992) Evolution of the vertebrate Hox homeobox genes. *Bioessays.* **14,** 245-252.

Krumlauf, R. (1993) Hox genes and pattern formation in the branchial region of the vertebrate head. *Trends Genet* **9,** 106-112.

Krumlauf, R. (1994) *Hox* genes in vertebrate development. *Cell* **78,** 191-201.

Kulesa, P., Bronner-Fraser, M., and Fraser, S. (2000) In ovo time-lapse analysis after dorsal neural tube ablation shows rerouting of chick hindbrain neural crest. *Development* **127,** 2843-2852.

Kulesa, P.M., and Fraser, S.E. (2000) In ovo time-lapse analysis of chick hindbrain neural crest cell migration shows cell interactions during migration to the branchial arches. *Development* **127,** 1161-72.

Lang, G., Gombert, W.M., and Gould, H.J. (2005) A transcriptional regulatory element in the coding sequence of the human Bcl-2 gene. *Immunology* **114,** 25-36.

Langston, A.W., and Gudas, L.J. (1992) Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene *Hox-1.6*. *Mech. Dev.* **38,** 217-228.

Lawrence, P.A., and Struhl, G. (1996) Morphogens, compartments, and pattern: lessons from drosophila? *Cell* **85,** 951-61.

Le Mouellic, H., Lallemand, Y., and Brulet, P. (1992) Homeosis in the mouse induced by a null mutation in the homeo-gene Hox-3.1. *Cell* **69,** 251-264.

Lee, P.N., Callaerts, P., De Couet, H.G., and Martindale, M.Q. (2003) Cephalopod Hox genes and the origin of morphological novelties. *Nature* **424,** 1061-5.

Leid, M., Kastner, P., and Chambon, P. (1992) Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends Biochem Sci* **17,** 427-433.

Lenhard, B., Sandelin, A., Mendoza, L., Engstrom, P., Jareborg, N., and Wasserman, W.W. (2003) Identification of conserved regulatory elements by comparative genome analysis. *J Biol* **2,** 13.

Levine, M., and Davidson, E.H. (2005) Gene regulatory networks for development. *Proc Natl Acad Sci U S A* **102,** 4936-42.

Lewis, E.B. (1978) A gene complex controlling segmentation in Drosophila. *Nature* **276,** 565-570.

Lufkin, T., Dierich, A., LeMeur, M., Mark, M., and Chambon, P. (1991) Disruption of the *Hox-1.6* homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell* **66,** 1105-1119.

Lumsden, A., and Keynes, R. (1989) Segmental patterns of neuronal development in the chick hindbrain. *Nature* **337,** 424-428.

Lumsden, A., Sprawson, N., and Graham, A. (1991) Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development* **113,** 1281-1291.

Lumsden, A., and Krumlauf, R. (1996) Patterning the vertebrate neuraxis. *Science* **274,** 1109-1115.

Lumsden, A. (2004) Segmentation and compartition in the early avian hindbrain. *Mech Dev* **121,** 1081-8.

Lutz, B., Lu, H.-C., Eichele, G., Miller, D., and Kaufman, T. (1996) Rescue of *Drosophila labial* null mutant by the chicken ortholog *Hoxb-1* demonstrates that the function of *Hox* genes is phylogenetically conserved. *Genes Dev.* **10,** 176-184.

Maconochie, M., Krishnamurthy, R., Nonchev, S., Meier, P., Manzanares, M., Mitchell, P.J., and Krumlauf, R. (1999) Regulation of *Hoxa2* in cranial neural crest cells involves members of the *AP-2* family. *Development* **126,** 1483-1494.

Maconochie, M.K., Nonchev, S., Studer, M., Chan, S.K., Popperl, H., Sham, M.H., Mann, R.S., and Krumlauf, R. (1997) Cross-regulation in the mouse *HoxB* complex: the expression of *Hoxb2* in rhombomere 4 is regulated by *Hoxb1*. *Genes and Development* **11,** 1885-1896.

Maconochie, M.K., Nonchev, S., Manzanares, M., Marshall, H., and Krumlauf, R. (2001) Differences in Krox20-dependent regulation of *Hoxa2* and *Hoxb2* during hindbrain development. *Dev. Biol.* **233,** 468-81.

Maden, M. (2002) Retinoid signalling in the development of the central nervous system. *Nat Rev Neurosci* **3,** 843-53.

Mainguy, G., In der Rieden, P.M., Berezikov, E., Woltering, J.M., Plasterk, R.H., and Durston, A.J. (2003) A position-dependent organisation of retinoid response elements is conserved in the vertebrate Hox clusters. *Trends Genet* **19,** 476-9.

Mangelsdorf, D.J., Ong, E.S., Dyck, J.A., and Evans, R.M. (1990) Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature* **345,** 224-229.

Manley, N.R., and Capecchi, M.R. (1997) *Hox* group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures. *Dev. Biol.* **192,** 274-288.

Mann, R., and Abu-Shaar, M. (1996) Nuclear import of the homeodomain protein extradenticle in response to Wg asnd Dpp signalling. *Nature* **383,** 630-633.

Mann, R., and Chan, S.-K. (1996) Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *TIG* **12,** 258-262.

Manzanares, M., Cordes, S., Kwan, C.-T., Sham, M.-H., Barsh, G., and Krumlauf, R. (1997) Segmental regulation of *Hoxb3* by *kreisler*. *Nature* **387,** 191-195.

Manzanares, M., Cordes, S., Ariza-McNaughton, L., Sadl, V., Maruthainar, K., Barsh, G., and Krumlauf, R. (1999a) Conserved and distinct roles of *kreisler* in regulation of the paralogous *Hoxa3* and *Hoxb3* genes. *Development* **126,** 759-769.

Manzanares, M., Trainor, P.A., Nonchev, S., Ariza-McNaughton, L., Brodie, J., Gould, A., Marshall, H., Morrison, A., Kwan, C.T., Sham, M.H., Wilkinson, D.G., and Krumlauf, R. (1999b) The role of *kreisler* in segmentation during hindbrain development. *Dev. Biol.* **211,** 220-237.

Manzanares, M., Bel-Vialer, S., Ariza-McNaughton, L., Ferretti, E., Marshall, H., Maconochie, M.K., Blasi, F., and Krumlauf, R. (2001) Independent regulation of initiation and maintenance phases of *Hoxa3* expression in the vertebrate hindbrain involves auto and cross-regulatory mechanisms. *Development* **128,** 3595-3607.

Manzanares, M., Nardelli, J., Gilardi-Hebenstreit, P., Marshall, H., Giudicelli, F., Martinez-Pastor, M.T., Krumlauf, R., and Charnay, P. (2002) Krox20 and kreisler cooperate in the transcriptional control of segmental expression of Hoxb3 in the developing hindbrain. *EMBO J.* **21,** 365-376.

Marin, F., and Puelles, L. (1995) Morphological fate of rhombomeres in quail/chick chimeras:a segmental analysis of hindbrain nuclei. *European J. Neuroscience* **7,** 1714-1738.

Marin, F., and Charnay, P. (2000) Hindbrain patterning: FGFs regulate Krox20 and mafB/kr expression in the otic/preotic region. *Development* **127,** 4925-35.

Markstein, M., Markstein, P., Markstein, V., and Levine, M.S. (2002) Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the Drosophila embryo. *Proc Natl Acad Sci U S A* **99,** 763-8.

Marshall, H., Nonchev, S., Sham, M.H., Muchamore, I., Lumsden, A., and Krumlauf, R. (1992) Retinoic acid alters hindbrain Hox code and induces transformation of rhombomeres 2/3 into a 4/5 identity. *Nature* **360,** 737-741.

Marshall, H., Studer, M., Pöpperl, H., Aparicio, S., Kuroiwa, A., Brenner, S., and Krumlauf, R. (1994) A conserved retinoic acid response element required for early expression of the homeobox gene *Hoxb-1*. *Nature* **370,** 567-571.

McClintock, J.M., Carlson, R., Mann, D.M., and Prince, V.E. (2001) Consequences of Hox gene duplication in the vertebrates: an investigation of the zebrafish Hox paralogue group 1 genes. *Development* **128,** 2471-84.

McClintock, J.M., Kheirbek, M.A., and Prince, V.E. (2002) Knockdown of duplicated zebrafish hoxb1 genes reveals distinct roles in hindbrain patterning and a novel mechanism of duplicate gene retention. *Development* **129,** 2339-54.

McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A., and Gehring, W. (1984) A conserved DNA sequence in homeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* **308,** 428-433.

McGinnis, W., and Krumlauf, R. (1992) Homeobox genes and axial patterning. *Cell* **68,** 283-302.

McKay, I.J., Muchamore, I., Krumlauf, R., Maden, M., Lumsden, A., and Lewis, J. (1994) The *kreisler* mouse: a hindbrain segmentation mutant that lacks two rhombomeres. *Development* **120,** 2199-2211.

Mechta-Grigoriou, F., Garel, S., and Charnay, P. (2000) Nab proteins mediate a negative feedback loop controlling Krox-20 activity in the developing hindbrain. *Development* **127,** 119-28.

Mendelsohn, C., Larkin, S., Mark, M., LeMeur, M., Clifford, J., Zelent, A., and Chambon, P. (1994a) RAR·β isoforms: distinct transcriptional control by retinoic acid and specific spatial patterns of promoter activity during mouse embryonic development. *Mech Dev* **45,** 227-241.

Mendelsohn, C., Lohnes, D., Decimo, D., Lufkin, T., LeMeur, M., Chambon, P., and Mark, M. (1994b) Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants. *Development* **120,** 2749-71.

Minguillon, C., Gardenyes, J., Serra, E., Castro, L.F., Hill-Force, A., Holland, P.W., Amemiya, C.T., and Garcia-Fernandez, J. (2005) No more than 14: the end of the amphioxus Hox cluster. *Int J Biol Sci* **1,** 19-23.

Misof, B.Y., and Wagner, G.P. (1996) Evidence for four Hox clusters in the killifish Fundulus heteroclitus (teleostei). *Mol Phylogenet Evol* **5,** 309-22.

Morgan, B.A., Izpisua-Belmonte, J.-C., Duboule, D., and Tabin, C.J. (1992) Targeted misexpression of Hox-4.6 in the avian limb bud causes apparent homeotic transformations. *Nature* **358,** 236-239.

Murakami, A., Thurlow, J., and Dickson, C. (1999) Retinoic acid-regulated expression of fibroblast growth factor 3 requires the interaction between a novel transcription factor and GATA-4. *J. Biol. Chem.* **274,** 17242-8.

Naruse, K., Fukamachi, S., Mitani, H., Kondo, M., Matsuoka, T., Kondo, S., Hanamura, N., Morita, Y., Hasegawa, K., Nishigaki, R., Shimada, A., Wada, H., Kusakabe, T., Suzuki, N., Kinoshita, M., Kanamori, A., Terado, T., Kimura, H., Nonaka, M., and Shima, A. (2000) A detailed linkage map of medaka, Oryzias latipes: comparative genomics and genome evolution. *Genetics* **154,** 1773-84.

Niederreither, K., Subbarayan, V., Dollé, P., and Chambon, P. (1999) Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* **21,** 444-448.

Niederreither, K., Vermot, J., Schuhbaur, B., Chambon, P., and Dollé, P. (2000) Retinoic acid synthesis and hindbrain patterning in the mouse embryo. *Development* **127,** 75-85.

Nieto, M.A., Gilardi-Hebenstreit, P., Charnay, P., and Wilkinson, D.G. (1992) A receptor protein tyrosine kinase implicated in the segmental patterning of the hindbrain and mesoderm. *Development* **116,** 1137-1150.

Noden, D. (1988) Interactions and fates of avian craniofacial mesenchyme. *In* "Development" (P. Thorogood and C. Tickle, Eds.), Vol. 103, pp. 121-140.

Noden, D.M. (1983) The role of the neural crest in patterning of avian cranial skeletal, connective, and muscle tissues. *Dev. Biol.* **96,** 144-165.

Nonchev, S., Maconochie, M., Vesque, C., Aparicio, S., Ariza-McNaughton, L., Manzanares, M., Maruthainar, K., Kuroiwa, A., Brenner, S., Charnay, P., and Krumlauf, R. (1996a) The conserved role of *Krox-20* in directing *Hox* gene expression during vertebrate hindbrain segmentation. *Proc. Natl. Acad. Sci. USA* **93,** 9339-9345.

Nonchev, S., Vesque, C., Maconochie, M., Seitanidou, T., Ariza-McNaughton, L., Frain, M., Marshall, H., Sham, M.H., Krumlauf, R., and Charnay, P. (1996b) Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20. *Development* **122,** 543-554.

Ohnemus, S., Bobola, N., Kanzler, B., and Mallo, M. (2001) Different levels of Hoxa2 are required for particular developmental processes. *Mech Dev* **108,** 135-47.

Ohno, S. (1993) Patterns in genome evolution. *Curr Opin Genet Dev* **3,** 911-4.

Otting, G., Qian, Y.Q., Muller, M., Affolter, M., Gehring, W., and Wuthrich, K. (1988) Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. *Junk* **7,** 4305-4309.

Oudejans, C.B., Pannese, M., Simeone, A., Meijer, C.J., and Boncinelli, E. (1990) The three most downstream genes of the Hox-3 cluster are expressed in human extraembryonic tissues including trophoblast of androgenetic origin. *Development.* **108,** 471-477.

Packer, A.I., Crotty, D.A., Elwell, V.A., and Wolgemuth, D.J. (1998) Expression of the murine Hoxa4 gene requires both autoregulation and a conserved retinoic acid response element. *Development* **125,** 1991-1998.

Papalopulu, N., Lovell-Badge, R., and Krumlauf, R. (1991) The expression of murine Hox-2 genes is dependent on the differentiation pathway and displays a collinear sensitivity to retinoic acid in F9 cells and Xenopus embryos. *Nuc. Acids Res.* **19,** 5497-5506.

Pasqualetti, M., Ori, M., Nardi, I., and Rijli, F.M. (2000) Ectopic Hoxa2 induction after neural crest migration results in homeosis of jaw elements in Xenopus. *Development* **127,** 5367-78.

Pata, I., Studer, M., van Doorninck, J.H., Briscoe, J., Kuuse, S., Engel, J.D., Grosveld, F., and Karis, A. (1999) The transcription factor GATA3 is a downstream effector of Hoxb1 specification in rhombomere 4. *Development* **126,** 5523-31.

Pennacchio, L.A., and Rubin, E.M. (2001) Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet* **2,** 100-9.

Percival-Smith, A., Weber, J., Gilfoyle, E., and Wilson, P. (1997) Genetic characterization of the role of the two HOX proteins, Proboscipedia and Sex Combs Reduced, in determination of adult antennal, tarsal, maxillary palp and proboscis identities in Drosophila melanogaster. *Development* **124,** 5049-62.

Percival-Smith, A., and Bondy, J.A. (1999) Analysis of murine HOXA-2 activity in Drosophila melanogaster. *Dev Genet* **24,** 336-44.

Piccolo, S., Sasai, Y., Lu, B., and De Robertis, E.M. (1996) Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* **86,** 589-98.

Pollock, R., Jay, G., and Bieberich, C. (1992) Altering the boundaries of *Hox-3.1* expression: evidence for antipodal gene regulation. *Cell* **71,** 911-924.

Pöpperl, H., and Featherstone, M. (1993) Identification of a retinoic acid repsonse element upstream of the murine *Hox-4.2* gene. *Mol. Cell. Biol.* **13,** 257-265.

Pöpperl, H., Bienz, M., Studer, M., Chan, S., Aparicio, S., Brenner, S., Mann, R., and Krumlauf, R. (1995) Segmental expression of *Hoxb1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/Pbx*. *Cell* **81,** 1031-1042.

Prince, V., and Lumsden, A. (1994) Hox-a2 expression in normal and transposed rhombomeres: independent regulation in the neural tube and neural crest. *Development* **120,** 911-923.

Prince, V.E., Joly, L., Ekker, M., and Ho, R.K. (1998) Zebrafish hox genes: genomic organization and modified colinear expression patterns in the trunk. *Development* **125,** 407-20.

Prince, V.E., and Pickett, F.B. (2002) Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet* **3,** 827-37.

Ramirez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R., and Bradley, A. (1993) Hoxb-4 (Hox-2.6) mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. *Cell* **73,** 279-294.

Randazzo, F., Cribbs, D., and Kaufman, T. (1991) Rescue and regulation of proboscipedia a homeotic gene of the Antennapedia Complex. *Development* **111,** 1-16.

Ren, S.Y., Angrand, P.O., and Rijli, F.M. (2002) Targeted insertion results in a Rhombomere 2-specific Hoxa2 knockdown and ectopic activation of Hoxa1 expression. *Dev Dyn* **225,** 305-15.

Rijli, F.M., Mark, M., Lakkaraju, S., Dierich, A., Dollé, P., and Chambon, P. (1993) A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* **75,** 1333-1349.

Rossel, M., and Capecchi, M.R. (1999) Mice mutant for both *Hoxa1* and *Hoxb1* show extensive remodeling of the hindbrain and defects in craniofacial development. *Development* **126,** 5027-5040.

Ruberte, E., Dollé, P., Chambon, P., and Morriss-Kay, G. (1991) Retinoic acid receptors and cellular retinoid binding proteins II. Their differential pattern of transcription during early morphogenesis in mouse embryos. *Development* **111,** 45-60.

Russo, M.W., Sevetson, B.R., and Milbrandt, J. (1995) Identification of NAB1, a repressor of NGFI-A- and Krox20-mediated transcription. *Proc Natl Acad Sci U S A* **92,** 6873-7.

Saint-Jeannet, J.P., He, X., Varmus, H.E., and Dawid, I.B. (1997) Regulation of dorsal fate in the neuraxis by Wnt-1 and Wnt-3a. *Proc Natl Acad Sci U S A* **94,** 13713-8.

Samad, O.A., Geisen, M.J., Caronia, G., Varlet, I., Zappavigna, V., Ericson, J., Goridis, C., and Rijli, F.M. (2004) Integration of anteroposterior and dorsoventral regulation of Phox2b transcription in cranial motoneuron progenitors by homeodomain proteins. *Development* **131,** 4071-83.

Scemama, J.L., Hunter, M., McCallum, J., Prince, V., and Stellwag, E. (2002) Evolutionary divergence of vertebrate Hoxb2 expression patterns and transcriptional regulatory loci. *J Exp Zool* **294,** 285-99.

Schilling, T. (2001) Plasticity of zebrafish Hox expression in the hindbrain and cranial neural crest hindbrain. *Dev. Biol.* **231,** 201-216.

Schilling, T.F., and Kimmel, C.B. (1994) Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development* **120,** 483-94.

Schneider, R.A., Hu, D., Rubenstein, J.L., Maden, M., and Helms, J.A. (2001) Local retinoid signaling coordinates forebrain and facial morphogenesis by maintaining FGF8 and SHH. *Development* **128,** 2755-67.

Schneider-Maunoury, S., Topilko, P., Seitandou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C., and Charnay, P. (1993) Disruption of *Krox-20* results in alteration of rhombomeres 3 and 5 in the developing hindbrain. *Cell* **75,** 1199-1214.

Schuler, G.D., Altschul, S.F., and Lipman, D.J. (1991) A workbench for multiple alignment construction and analysis. *Proteins* **9,** 180-90.

Sechrist, J., and Bronner-Fraser, M. (1991) Birth and differentiation of reticular neurons in the chick hindbrain: ontogeny of the first neuronal population. *Neuron* **7,** 947-963.

Sechrist, J., Serbedzija, G.N., Scherson, T., Fraser, S.E., and Bronner-Fraser, M. (1993) Segmental migration of the hindbrain neural crest does not arise from its segmental generation. *Development* **118(3),** 691-703.

Seitanidou, T., Schneider-Maunoury, S., Desmarquet, C., Wilkinson, D., and Charnay, P. (1997) *Krox20* is a key regulator of rhombomere-specific gene expression in the developing hindbrain. *Mech. Dev.* **65,** 31-42.

Sham, M.H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Das Gupta, R.D., Whiting, J., Wilkinson, D., Charnay, P., and Krumlauf, R. (1993) The zinc finger gene *Krox-20* regulates *Hoxb-2 (Hox2.8)* during hindbrain segmentation. *Cell* **72,** 183-196.

Sharpe, J., Nonchev, S., Gould, A., Whiting, J., and Krumlauf, R. (1998) Selectivity, sharing and competitive interactions in the regulation of *Hoxb* genes. *EMBO J.* **17,** 1788-1798.

Simeone, A., Acampora, D., D'Esposito, M., Faiella, A., Pannese, M., Scotto, L., Montanucci, M., D'Alessandro, G., Mavilio, F., and Boncinelli, E. (1989) Posttranscriptional control of human homeobox gene expression in induced NTERA-2 embryonal carcinoma cells. *Mol. Reprod. Dev.* **1,** 107-115.

Snell, E.A., Scemama, J.L., and Stellwag, E.J. (1999) Genomic organization of the Hoxa4-Hoxa10 region from Morone saxatilis: implications for Hox gene evolution among vertebrates. *J Exp Zool* **285,** 41-9.

Sockanathan, S., and Jessell, T.M. (1998) Motor neuron-derived retinoid signaling specifies the subtype identity of spinal motor neurons. *Cell* **94,** 503-14.

Spemann, H., and Mangold, H. (2001) Induction of embryonic primordia by implantation of organizers from a different species. 1923. *Int J Dev Biol* **45,** 13-38.

Spitz, F., Gonzalez, F., and Duboule, D. (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113,** 405-17.

Stormo, G.D., and Fields, D.S. (1998) Specificity, free energy and information content in protein-DNA interactions. *Trends Biochem Sci* **23,** 109-13.

Stormo, G.D. (2000) Identification of coordinated gene expression and regulatory sequences. *Pac Symp Biocomput,* 416-7.

Studer, M., Pöpperl, H., Marshall, H., Kuroiwa, A., and Krumlauf, R. (1994) Role of a conserved retinoic acid response element in rhombomere restriction of *Hoxb-1. Science* **265,** 1728-1732.

Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A., and Krumlauf, R. (1996) Altered segmental identity and abnormal migration of motor neurons in mice lacking *Hoxb-1. Nature* **384,** 630-635.

Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F., Chambon, P., and Krumlauf, R. (1998) Genetic interaction between *Hoxa1* and *Hoxb1* reveal new roles in regulation of early hindbrain patterning. *Development* **125,** 1025-1036.

Sucov, H.M., Murakami, K.K., and Evans, R.M. (1990) Characterization of an autoregulated response element in the mouse retinoic acid receptor type beta gene. *Proc Natl Acad Sci U S A* **87,** 5392-6.

Sucov, H.M., Dyson, E., Gumeringer, C.L., Price, J., Chien, K.R., and Evans, R.M. (1994) RXRα mutant mice establish a genetic basis for vitamin A signalling in heart morphogenesis. *Genes Dev.* **8,** 1007-1018.

Sucov, H.M., Izpisua Belmonte, J.C., Ganan, Y., and Evans, R.M. (1995) Mouse embryos lacking RXR alpha are resistant to retinoic-acid-induced limb defects. *Development* **121,** 3997-4003.

Svaren, J., Sevetson, B.R., Apel, E.D., Zimonjic, D.B., Popescu, N.C., and Milbrandt, J. (1996) NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by proliferative and differentiative stimuli. *Mol. Cell. Biol.* **16,** 3545-53.

Swiatek, P.J., and Gridley, T. (1993) Perinatal lethality and defects in hindbrain development in mice homozygous for a targeted mutation of the zinc finger gene *Krox-20. Genes Dev.* **7,** 2071-2084.

Swirnoff, A.H., Apel, E.D., Svaren, J., Sevetson, B.R., Zimonjic, D.B., Popescu, N.C., and Milbrandt, J. (1998) Nab1, a corepressor of NGFI-A (Egr-1), contains an active transcriptional repression domain. *Mol. Cell. Biol.* **18,** 512-24.

Takio, Y., Pasqualetti, M., Kuraku, S., Hirano, S., Rijli, F.M., and Kuratani, S. (2004) Evolutionary biology: lamprey Hox genes and the evolution of jaws. *Nature* **429,** 1 p following 262.

Theil, T., Frain, M., Gilardi-Hebenstreit, P., Flenniken, A., Charnay, P., and Wilkinson, D.G. (1998) Segmental expression of the EphA4 (Sek-1) receptor tyrosine kinase in the hindbrain is under direct transcriptional control of Krox-20. *Development* **125,** 443-52.

Theil, T., Ariza-McNaughton, L., Manzanares, M., Brodie, J., Krumlauf, R., and Wilkinson, D.G. (2002) Requirement for downregulation of kreisler during late patterning of the hindbrain. *Development* **129,** 1477-85.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nuc. Acids Res.* **22,** 4673-4680.

Thompson, J.R., Chen, S.W., Ho, L., Langston, A.W., and Gudas, L.J. (1998) An evolutionary conserved element is essential for somite and adjacent mesenchymal expression of the Hoxa1 gene. *Dev Dyn* **211,** 97-108.

Tickle, C., Alberts, B., Wolpert, L., and Lee, J. (1982) Local application of retinoic acid to the limb bud mimics the action of the polarizing region. *Nature* **296,** 564-566.

Toulouse, A., Morin, J., Pelletier, M., and Bradley, W.E. (1996) Structure of the human retinoic acid receptor beta 1 gene. *Biochim Biophys Acta* **1309,** 1-4.

Trainor, P., and Krumlauf, R. (2000) Plasticity in mouse neural crest cells reveals a new patterning role for cranial mesoderm. *Nature Cell Biology* **2,** 96-102.

Trainor, P.A., Tan, S.S., and Tam, P.P. (1994) Cranial paraxial mesoderm: regionalisation of cell fate and impact on craniofacial development in mouse embryos. *Development* **120,** 2397-408.

Trainor, P.A., Ariza-McNaughton, L., and Krumlauf, R. (2002) Role of the isthmus and FGFs in resolving theparadox of neural crest plasticity and prepatterning. *Science* **295,** 1288-91.

Tümpel, S., Maconochie, M., Wiedemann, L.M., and Krumlauf, R. (2002) Conservation and diversity in the cis-regulatory networks that integrate information controlling expression of Hoxa2 in hindbrain and cranial neural crest cells in vertebrates. *Dev. Biol.* **246,** 45-56.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., et al. (2001) The sequence of the human genome. *Science* **291,** 1304-51.

Vesque, C., Maconochie, M., Nonchev, S., Ariza-McNaughton, L., Kuroiwa, A., Charnay, P., and Krumlauf, R. (1996) Hoxb-2 transcriptional activation in rhombomeres 3 and 5 requires an evolutionarily conserved cis-acting element in addition to the Krox-20 binding site. *EMBO J.* **15,** 5383-5896.

Walshe, J., Maroon, H., McGonnell, I.M., Dickson, C., and Mason, I. (2002) Establishment of hindbrain segmental identity requires signaling by FGF3 and FGF8. *Curr Biol* **12,** 1117-23.

Weatherbee, S.D., Halder, G., Kim, J., Hudson, A., and Carroll, S. (1998) Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the Drosophila haltere. *Genes Dev.* **12,** 1474-82.

Weaver, D.C., Workman, C.T., and Stormo, G.D. (1999) Modeling regulatory networks with weight matrices. *Pac Symp Biocomput*, 112-23.

Wendling, O., Ghyselinck, N.B., Chambon, P., and Mark, M. (2001) Roles of retinoic acid receptors in early embryonic morphogenesis and hindbrain patterning. *Development* **128,** 2031-8.

Wiellette, E.L., and Sive, H. (2003) vhnf1 and Fgf signals synergize to specify rhombomere identity in the zebrafish hindbrain. *Development* **130,** 3821-9.

Wilkinson, D.G., Bhatt, S., Cook, M., Boncinelli, E., and Krumlauf, R. (1989) Segmental expression of Hox-2 homeobox-containing genes in the developing mouse hindbrain. *Nature* **341,** 405-409.

Wizenmann, A., and Lumsden, A. (1997) Segregation of rhombomeres by differential chemoaffinity. *Molecular and Cellular Neuroscience* **9,** 448-459.

Workman, C.T., and Stormo, G.D. (2000) ANN-Spec: a method for discovering transcription factor binding sites with improved specificity. *Pac Symp Biocomput*, 467-78.

Xu, Q., Alldus, G., Holder, N., and Wilkinson, D.G. (1995) Expression of truncated *Sek-1* receptor tyrosine kinase disrupts the segmental restriction of gene expression in the *Xenopus* and zebrafish hindbrain. *Development* **121,** 4005-4016.

Xu, Q., Mellitzer, G., Robinson, V., and Wilkinson, D.G. (1999) *In vivo* cell sorting in complementary segmental domains mediated by *Eph* receptors and *ephrins*. *Nature* **399,** 267-271.

Xu, Q., Mellitzer, G., and Wilkinson, D.G. (2000) Roles of Eph receptors and ephrins in segmental patterning. *Philos Trans R Soc Lond B Biol Sci* **355,** 993-1002.

Yee, S.-P., and Rigby, P.W.J. (1993) The regulation of *myogenin* gene expression during the embryonic development of the mouse. *Genes Dev.* **7,** 1277-1289.

Zakany, J., Gérard, M., Favier, B., and Duboule, D. (1997) Deletion of a *HoxD* enhancer induces transcriptional heterochrony leading to transposition of the sacrum. *EMBO J.* **16,** 4393-4402.

Zakany, J., Kmita, M., Alarcon, P., de la Pompa, J., and Duboule, D. (2001) Localized and Transient Transcription of Hox Genes Suggests a Link between Patterning and the Segmentation Clock. *Cell* **106,** 207-17.

Zelent, A., Mendelsohn, C., Kastner, P., Krust, A., Garnier, J.M., Ruffenach, F., Leroy, P., and Chambon, P. (1991) Differentially expressed isoforms of the mouse retinoic acid receptor beta generated by usage of two promoters and alternative splicing. *EMBO J.* **10,** 71-81.

Zhang, F., Nagy Kovacs, E., and Featherstone, M.S. (2000) Murine hoxd4 expression in the CNS requires multiple elements including a retinoic acid response element. *Mech Dev* **96,** 79-89.

Zhang, J., and Nei, M. (1996) Evolution of Antennapedia-class homeobox genes. *Genetics* **142,** 295-303.

Zhang, M., Kim, H.J., Marshall, H., Gendron-Maguire, M., Lucas, D.A., Baron, A., Gudas, L.J., Gridley, T., Krumlauf, R., and Grippo, J.F. (1994) Ectopic *Hoxa-1* induces rhombomere transformation in mouse hindbrain. *Development* **120,** 2431-2442.

Zimmerman, L.B., De Jesús-Escobar, J.M., and Harland, R.M. (1996) The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* **86,** 599-606.